

RICE UNIVERSITY

The Role of Executive Control in Garden Path Reinterpretation

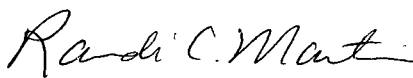
by

Loan Cam Vuong

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:



Randi C. Martin, Elma Schneider Professor, Chair,
Psychology



David M. Lane, Associate Professor,
Psychology, Statistics, and Management



Jessica Logan, Assistant Professor,
Psychology



Suzanne Kemmer, Associate Professor,
Linguistics

HOUSTON, TEXAS
MAY 2010

UMI Number: 3421435

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

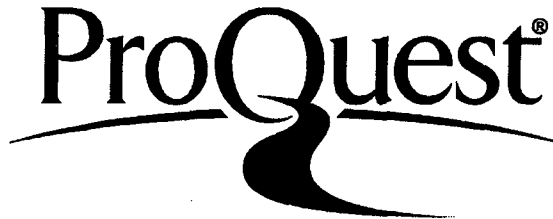
In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3421435

Copyright 2010 by ProQuest LLC.

All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

The Role of Executive Control in Garden Path Reinterpretation

by

Loan Cam Vuong

Language comprehension is a complex process in which multiple, including semantic and syntactic, representations of the material are constructed as the input unfolds. Sometimes more than one way to interpret a sentence or sequence of sentences is possible and reliable sources of information may mislead us into developing an ultimately incorrect interpretation. Several questions could be asked about language comprehension in these cases. First, how good are we at recovering from misinterpretations? Second, do we vary in ability to recover from misinterpretations? Third, if there are individual differences in recovery ability, what cognitive factors underlie the differences? And fourth, are those cognitive factors subserved by specific areas in the brain?

Answers to the first question may depend on the type of sentences involved (Gorrell, 1995; Pritchett, 1992; Sturt, 1995; Sturt, Pickering, & Crocker, 1999). Some sentences appear to be hard (e.g., *The horse raced past the barn fell*, cf. the horse that was raced past the barn fell) (Bever, 1970) while others are relatively easy to reinterpret (e.g., *John saw the girl was cheating*, Ferreira & Henderson, 1990). Several theories (e.g., Gorrell, 1995; Pritchett, 1992; Sturt, 1995) categorize sentences into those that could be recovered “unconsciously” – the ones we typically do well at reinterpreting and those that are “consciously” difficult, which we may or may not succeed in reinterpreting.

Using sentences shown to induce noticeable degree of reinterpretation difficulty (the “garden path” sentences), this thesis examined the role of executive control,

particularly ability to attend to relevant information and ignore salient but irrelevant information, in reinterpretation. To the extent that executive control ability varies from individual to individual, ability to recover from misinterpretations should vary accordingly. Further, the brain areas that are important for executive control, particularly the left inferior frontal gyrus (LIFG), should be important for recovering from misinterpretations. Experiment 1-3 examined the issue by focusing on individual differences in executive control among healthy younger comprehenders. Experiment 4-6 examined the same issue by focusing on LIFG-based executive control patients. The thesis shows that LIFG-based executive control is critical for both semantic and syntactic reinterpretation of garden path sentences.

Acknowledgements

This work was supported by NIH grant no. DC-00128 to Randi Martin and an APA Dissertation Research Award to myself. I would like to thank Catherine Johnston, Elaine Mahoney, and Aurora Ramos for assistance with data collection. Thanks also to Cris Hamilton for sharing test materials used in Experiment 4. I am grateful to my committee members, David Lane, Jessica Logan, Randi Martin, and Suzanne Kemmer for very helpful questions and suggestions. And of course, I can never thank Randi Martin enough for her time, patience, and guidance; she is a wonderful advisor.

I am also very fortunate to have had the chance to study with Dr. Randy Diehl, Arthur Markman, and David Birdsong – my senior thesis advisor – in college.

Special thanks to my family, especially my parents, for their incredible generosity and support.

Table of Contents

Chapter 1	Introduction	1
Chapter 2	Individual Differences in Executive Control and Reinterpretation	10
	Experiment 1. Executive Control and Semantic Recovery	10
	Experiment 2. Garden Path Sentence Comprehension: A Follow-up	35
	Experiment 3. Executive Control and Syntactic Recovery	42
Chapter 3	Executive Control Deficit and Reinterpretation	56
	Patient Background	56
	Experiment 4. LIFG Damage and Executive Control	63
	Experiment 5. LIFG Damage and Semantic Recovery	70
	Experiment 6. LIFG Damage and Syntactic Recovery	90
Chapter 4	Summary of Findings and General Discussion	106
	References	113
	Appendix A	122
	Appendix B	124
	Appendix C	125

List of Tables

Table 1.1	Examples of materials for garden path sentence processing tasks in Experiment 1-3.	17
Table 1.2	Properties of garden path sentence materials used in Experiment 1-3 and 5-6.	19
Table 1.3	Experiment 1 – garden path comprehension among younger comprehenders: Mean accuracy (and standard errors) by condition and <i>F</i> -statistics.	23
Table 1.4	Experiment 1 – garden path comprehension among younger comprehenders: Mean residual reading times (and standard errors) by condition at the main verb region and <i>F</i> -statistics.	25
Table 1.5	Experiment 1 – garden path comprehension among younger comprehenders: Mean residual reading times (and standard errors) by condition at the final region and <i>F</i> -statistics.	27
Table 1.6a	Experiment 1 – correlation across younger comprehenders: Means, standard deviations, and reliability estimates for RT measures (final region).	29
Table 1.6b	Experiment 1 – correlation across younger comprehenders: Means, standard deviations, and reliability estimates for accuracy measures.	29
Table 1.7	Experiment 1 – correlation across younger comprehenders:	30

	Relationship between executive control and semantic recovery.	
Table 2.1	Experiment 2 – garden path comprehension among younger comprehenders: Mean accuracy (and standard errors) by condition and <i>F</i> -statistics.	38
Table 2.2	Experiment 2 – garden path comprehension among younger comprehenders: Mean residual reading times (and standard errors) by condition at the main verb region and <i>F</i> -statistics.	39
Table 2.3	Experiment 2 – garden path comprehension among younger comprehenders: Mean residual reading times (and standard errors) by condition at the final region and <i>F</i> -statistics.	40
Table 2.4	Experiment 2 – garden path comprehension among younger comprehenders: Mean untrimmed reading times by condition for all trials and for correct vs. incorrect trials.	41
Table 3.1	Experiment 3 – garden path grammaticality judgment among younger comprehenders: Mean accuracy (and standard errors) by condition and <i>F</i> -statistics.	48
Table 3.2	Experiment 3 – garden path grammaticality judgment among younger comprehenders: Mean residual reading times (and standard errors) by condition at the main verb region and <i>F</i> -statistics.	49
Table 3.3	Experiment 3 – garden path grammaticality judgment among	50

	younger comprehenders: Mean residual reading times (and standard errors) by condition at the final region and <i>F</i> -statistics.	
Table 3.4	Experiment 3 – garden path grammaticality judgment among younger comprehenders: Mean untrimmed reading times by condition for all trials and for correct vs. incorrect trials.	51
Table 3.5a	Experiment 3 – correlation across younger comprehenders: Means, standard deviations, and reliability estimates for accuracy measures.	52
Table 3.5b	Experiment 3 – correlation across younger comprehenders: Means, standard deviations, and reliability estimates for RT measures.	52
Table 3.6	Experiment 3 – correlation across younger comprehenders: Relationship between executive control and syntactic recovery.	53
Table 4.1	Background of the patients tested in Experiment 4-6.	58
Table 4.2	Experiment 4 – executive control: Mean interference effects (and standard deviations) for normal controls and patients.	66
Table 5.1a	Experiment 5 – executive control: Mean verbal Stroop effects (and standard deviations) for normal controls in Experiment 5-6 vs. normal controls in Experiment 4.	74

Table 5.1b	Experiment 5 – executive control: Mean verbal Stroop effects (and standard errors) by condition for younger comprehenders in Experiment 1 vs. for normal (older) controls in Experiment 5-6.	75
Table 5.2	Experiment 5 – overall sentence comprehension: Mean accuracy (and standard deviations) on experimental and filler sentences for normal controls and patients.	77
Table 5.3a	Experiment 5 – garden path comprehension: Means (and standard errors) by condition for normal controls.	78
Table 5.3b	Experiment 5 – garden path comprehension: <i>F</i> -statistics for normal controls.	78
Table 5.4	Experiment 5 – garden path comprehension: Mean accuracy (and standard errors) by condition for normal controls and patients, and <i>F</i> -statistics from item analysis for patients.	81
Table 5.5a	Experiment 5 – garden path comprehension: Mean reading times (and standard errors) by condition at the main verb region for normal controls and patients, and <i>F</i> -statistics from item analysis for patients.	85
Table 5.5b	Experiment 5 – garden path comprehension: Mean reading times (and standard errors) by condition at the final region for normal controls and patients, and <i>F</i> -statistics from item	86

analysis for patients.

Table 6.1	Experiment 6 –overall grammaticality judgment: Mean accuracy (and standard deviations) on experimental and filler sentences for normal controls and patients.	94
Table 6.2a	Experiment 6 – garden path grammaticality judgment: Means (and standard errors) by condition for normal controls.	95
Table 6.2b	Experiment 6 – garden path grammaticality judgment: <i>F</i> -statistics for normal controls.	95
Table 6.3	Experiment 6 – garden path grammaticality judgment: Mean accuracy (and standard errors) by condition for normal controls and patients, and <i>F</i> -statistics from item analysis for patients.	97
Table 6.4	Experiment 6 – garden path grammaticality judgment: Mean reading times (and standard errors) by condition at the main verb region for normal controls and patients, and <i>F</i> -statistics from item analysis for patients.	100
Table 6.5a	Experiment 6 – garden path grammaticality judgment: Mean accuracy for younger comprehenders in Experiment 3 (implausible conditions only), for returning vs. fresh group of normal (older) controls, and for patients.	103
Table 6.5b	Experiment 6 – garden path grammaticality judgment: Mean accuracy (and standard errors) by condition and <i>F</i> -statistics	103

for fresh group of normal controls.

List of Figures

Figure 1a	Experiment 1: Relationship between verbal Stroop effect and ambiguity effect on reading time in the plausible, equi-biased verb condition across participants.	32
Figure 1b	Experiment 1: Relationship between verbal Stroop effect and ambiguity effect on reading time in the plausible, transitive-biased verb condition across participants.	32
Figure 2	Experiment 3: Relationship between verbal Stroop effect and overall ambiguity effect on grammaticality judgment accuracy across participants.	54
Figure 3	Structural brain scan of LIFG patient ML.	61
Figure 4	Structural brain scan of LIFG patient DW.	62

Chapter 1. Introduction

1.1. Garden path sentences

Ambiguity is pervasive in human language, though we may not notice it. When someone says, “She talked to the bank yesterday,” we interpret *bank* as referring to a financial institution immediately without much thinking although the ambiguous word *bank* could potentially refer to a completely different entity – a river’s edge. However, there are times when effects of ambiguity turn more apparent, as, for example, when someone says:

The old man’s glasses are filled with sherry. (Lewis, 1993)

The ambiguity here regarding the meaning of *glasses* (eye glasses vs. drinking glasses) is likely to be more noticeably felt. Similar situations occur beyond the word level. While the ambiguity in *She saw the man with the binoculars* is hardly obvious to most of us, in which the prepositional phrase *with the binoculars* could attach to *saw* vs. *the man* (i.e., using the binoculars she saw the man vs. she saw the man who was in possession of the binoculars), the ambiguity in

I convinced her professors hate me. (Pritchett, 1988)

causes more thinking, the outcome of which may or may not be satisfactory (the correct reading has *professors* attached as the subject of the complement clause – I convinced her that professors hate me).

Garden path sentences are a subset of the sentences containing an ambiguity. In garden path sentences, the wrong choice is made at the point of a local ambiguity, resulting in an ultimately incorrect interpretation that must be revised for a correct

understanding of the sentence. In *syntactic* garden paths, more than one structure could be assigned to the material but the wrong structure is initially pursued and later elements could not be integrated into the current structure without breaking grammatical rules.

Besides the previous example, consider the following sentence:

While Anna bathed the baby spit up on the bed. (Christianson et al., 2001)

In this sentence, the subordinate clause contains the ambiguous verb *bathed*, which could be transitive (i.e., needs a direct object) or intransitive (i.e., needs no object). When the transitive structure is assigned to the verb, the noun phrase *the baby* attaches as the object of *bathed*. The result of pursuing the transitive structure is that the verb in the main clause *spit* could not be grammatically integrated into the structure, as no noun phrases remain available in the subject position for a legitimate attachment of *spit*.

In the above example, the verb in the main clause (*spit*) provides a syntactic error cue that signals that the pursued structure is wrong. Comprehenders typically experience a *garden path effect* at the syntactic error cue, reading that region of the sentence significantly slower than reading the same region of a corresponding unambiguous sentence (e.g., *While Anna bathed, the baby spit up on the bed*). Garden path sentences may also induce an increase in comprehension errors, with comprehenders holding on to the initial incorrect interpretation (e.g., Anna bathed the baby), failing to arrive at the ultimately correct interpretation (e.g., While Anna bathed herself the baby spit up on the bed) (e.g., Christianson et al., 2001; Patson et al., 2009).

Use of syntactic garden path sentences in paradigms that are sensitive to moment-by-moment sentence decoding processes, along with detailed manipulations of characteristics of test materials, has allowed psycholinguists to discover linguistic factors

(e.g., syntax, semantics, referential and discourse context) that partake in earliest stages of sentence processing. However, syntactic garden path sentences (henceforth garden path sentences) are of interest to this work for another reason, that is they induce misinterpretations to which comprehenders are typically strongly committed. The revision of these misinterpretations is often challenging, thus providing a good opportunity for examining the contribution of controlled mechanisms to language understanding.

1.2. The case for a role of executive control in garden path reinterpretation

For garden path sentences, the interpretation that is chosen at the point of ambiguity is not the result of a haphazard choice. Rather, it is the preferred interpretation of that segment of the material. The preference may originate from general parsing heuristics that prefer certain structures to others (e.g., Frazier & Fodor, 1978), from word biases (e.g., Boland & Cutler, 1996; MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994) or from general biases for certain structures based on prior exposure with the language (e.g., Mitchell, Cuetos Corley, & Brysbaert, 1995). Other aspects of the material, such as semantic and discourse context, may also jointly point to the ultimately wrong interpretation. For example, with an initial string like *While the mother cooked the chicken ...*, most comprehenders would interpret *the chicken* as the object of *cooked*. That object interpretation is strongly preferred because, among other factors, the verb *cook* prefers to occur with an object and *the chicken* fits well as something that can be cooked.

When an error cue appears and the new material cannot be grammatically integrated into the pursued interpretation, in searching for an alternative interpretation of the sentence, the comprehender must give priority to the new cue(s) and ignore the prior cues that mismatch the current cues. Otherwise, the same incorrect interpretation would be maintained, as the prior misleading cues still support the incorrect interpretation. However, as mentioned above, the incorrect interpretation is the system's preferred interpretation with the prior cues tending to be highly reliable in normal circumstances. Given their saliency, these cues could strongly interfere in the reinterpretation process. Ability to resolve interference from irrelevant or no-longer-relevant information should, therefore, be important for successful garden path reinterpretation.

Generally, tasks that involve resolving interference from irrelevant information are thought to involve central executive control. The central executive is a theoretical construct that is postulated as a component of working memory – a system that temporarily stores and manipulates information (Baddeley & Hitch, 1994). The central executive is assumed to serve attentional control functions, such as focusing attention under condition of interference from irrelevant information, switching attention between different sources of information, and dividing attention during multi-tasking (Baddeley, 1986, 1998; Baddeley & Logie, 1999). A classic example of an executive control task is the verbal Stroop task, which involves naming the ink colors of written stimuli (Stroop, 1935). In one condition of the task, the stimuli are nonwords printed in color ink (e.g., rows of asterisks printed in *blue*). In another condition, the stimuli are color words printed in a mismatching ink color (e.g., the word *red* printed in *blue* ink). In the mismatch

condition, participants need to ignore the automatically read words and focus on the colors of the ink to generate correct responses. Compared to the nonword condition, ink color naming in the mismatch condition may be delayed by over 100 ms and more likely to produce naming errors (e.g., Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000, see MacLeod, 1991 for a review).

Given the functional requirements hypothesized to be necessary for garden path reinterpretation (e.g., resolving interference from the prepotent garden path interpretation during reinterpretation), it is parsimonious to assume that the mechanisms responsible for coordinating the garden path reinterpretation process are the same mechanisms served by the central executive (see Novick et al., 2005 for a review). Given that hypothesis, several predictions could be made on issues of individual differences and brain localization for garden path recovery. First, given that there are individual differences in executive control ability (e.g., Miyake et al., 2000), there should be individual differences in garden path recovery ability. In other words, not all native language comprehenders are equal in their ability to recover from garden path misinterpretations. Individuals who have better executive control should be better at garden path recovery. Further, given that the left inferior frontal gyrus (LIFG) area in the brain is important for executive control (e.g., Hamilton & Martin, 2005, 2007; Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998; Milham, Banich, & Baah, 2003; Thompson-Schill et al., 2002; see Jonides & Nee, 2006 for a review), the LIFG should be important for garden path recovery.

A few studies have provided encouraging evidence for a role of the LIFG in garden path reinterpretation (January, Trueewswell, & Thompson-Schill, 2009; Mason,

Just, Keller, & Carpenter, 2003; Novick, Trueswell, & Thompson-Schill, 2009, see also Novick et al., 2005 for a summary of preliminary evidence). In Mason et al. (Experiment 1), brain activity of neurologically healthy participants was recorded in a functional magnetic resonance imaging scanner while they read garden path or ambiguous control sentences (e.g., *The experienced soldiers warned about the dangers conducted the midnight raid* vs. *The experienced soldiers warned about the dangers before the midnight raid*, respectively). Consistent with the hypothesis that the LIFG is involved in garden path reinterpretation, brain activity in the LIFG significantly increased when participants were reading the garden path sentences compared to when they were reading the control sentences. Stronger evidence showing that the LIFG is necessary for garden path reinterpretation was obtained in a recent study by Novick et al. (2009, Experiment 4), who tested a stroke patient with LIFG damage, two control patients without LIFG damage, and a group of healthy control participants on the comprehension of garden path and unambiguous control sentences such as *Put the frog on the napkin in the box* vs. *Put the frog that's on the napkin in the box*, respectively. While the controls and non-LIFG patients were able to reinterpret the first prepositional phrase *on the napkin* as modifying the object noun phrase *the frog* and the second prepositional phrase *in the box* as the destination for the movement (in the box is where the frog on the napkin should be placed) without much problem (92% and 100% correct, respectively), the LIFG patient failed to do so on more than half of the garden path trials (38% correct). Importantly, the LIFG patient did not have a general sentence comprehension deficit, as shown by his perfect accuracy on control sentences.

Taken together, the evidence suggests that the LIFG is critical for garden path reinterpretation, which in turn provides support for the hypothesis that central executive control plays an important part in the garden path reinterpretation process. There is also some, albeit not strong, individual differences evidence among younger healthy comprehenders that supports the executive control hypothesis. In one study (Mendelsohn, 2002, Experiment 3), variation in the size of an executive control measure (obtained from a modified version of the Wisconsin Card Sorting Test constructed by the author) was found to correlate positively with variation in the size of a garden path effect across participants. However, the results were only suggestive because the garden path effect was manifested in the time taken to read comprehension questions and respond to questions (e.g., sentence: *Bill knew (that) the truth was being kept from him*, question: *Did Bill know the truth?*). The relationship between executive control and garden path reinterpretation that involves more direct measures of garden path reinterpretation, such as comprehension accuracy or time taken to reinterpret the garden path sentence during actual sentence reading, remains to be examined.

1.3. Further issues and the goal of the thesis

During sentence processing, different representations, including syntactic (i.e., grammatical) and semantic (i.e., meaning) representations of the material, are derived. Each of these representations will presumably need to be adjusted during the reinterpretation process. Logically, adjustment at one level of representation could be independent from adjustment at another level. A recent study suggests that it is indeed the case (Sturt, 2007). In that study, ambiguous (object/sentential complement garden path) and unambiguous control sentences were used (see below). The critical manipulation

involved varying the semantic consistency between later material and the initial misinterpretation.

Consistent

Ambiguous: The explorers found the South Pole was actually right at their feet.

Unambiguous: The explorers found *that* the South Pole was actually right at their feet.

Inconsistent

Ambiguous: The explorers found the South Pole was actually impossible to reach.

Unambiguous: The explorers found *that* the South Pole was actually impossible to reach.

For the ambiguous sentences, comprehenders tend to take the noun phrase following the first verb as the object of the verb (“the explorers found the South Pole”). The second verb (*was*), then, serves as a syntactic error cue that signals the object interpretation of the noun phrase is wrong (i.e., *was* needs a noun phrase to function as its grammatical subject but none is available, as *the explorers* has been taken as the subject and *the South Pole* has been taken as the object of *found*). Note that although the syntactic error cue is present in both the semantic conditions, semantic evidence contradicting the initial interpretation is available in the semantically inconsistent condition only (initial interpretation: the explorers found the South Pole vs. later semantic information: the South Pole could not be reached). Two garden path effects were found in Sturt’s (2007) study. The first effect was found at the second verb region, where the syntactic error cue was located, and was present in both the semantic conditions. The second effect was found at the final region, where the semantic error cue was located, and was specific to the semantically inconsistent condition. The first effect presumably

reflects syntactic reinterpretation whereas the second effect likely reflects semantic reinterpretation.

If garden path reinterpretation occurs in an independent fashion across different levels of representation, questions arise regarding the role of executive control in garden path reinterpretation at each representational level. In considering syntactic reinterpretation (i.e., reinterpreting the grammatical relations of elements, e.g., subject vs. object of verbs, in the sentence) vs. semantic reinterpretation (i.e., reinterpreting the thematic relations of elements, e.g., agent vs. patient of actions, in the sentence), it is logically possible that syntactic reinterpretation is carried out automatically while semantic reinterpretation involves executive control, or syntactic reinterpretation involves executive control while semantic reinterpretation is carried out automatically, or both of the processes are carried out automatically, or both involve executive control. The answers to this issue are still open, as it has not been empirically addressed in the literature.

The central goal of this thesis was to examine the role of executive control separately for semantic and syntactic garden path reinterpretation. Experiment 1-3 examined the issue by focusing on individual differences in executive control and garden path reinterpretation among healthy younger comprehenders. Experiment 4-6 examined the same issue by focusing on deficit in executive control and garden path reinterpretation in LIFG-damaged patients.

Chapter 2.

Individual Differences in Executive control and Garden Path Reinterpretation

Two individual differences experiments were carried out to examine the relationship between executive control and garden path reinterpretation at the semantic (Experiment 1) and syntactic level (Experiment 3). Participants performed the verbal Stroop task (Stroop, 1935), a nonverbal version of the Stroop task, and a garden path processing task in each experiment. The verbal Stroop task and the nonverbal Stroop task measure executive control in the verbal and nonverbal domain respectively. Prior evidence suggests that executive control ability may be separable for verbal vs. nonverbal materials (Hamilton & Martin, 2005). Both versions of the Stroop task were included to address this issue of domain specificity in executive control as it relates to garden path recovery. In addition to Experiment 1 and 3, another experiment (Experiment 2) was carried out that included only a garden path comprehension task. Experiment 2 was run to address issues that arose in Experiment 1 about garden path sentence comprehension.

Experiment 1. Executive Control and Semantic Recovery

Experiment 1 focused on semantic reinterpretation of garden path sentences. Participants read ambiguous (object/subject garden path) and unambiguous control sentences (see examples) one phrase at a time at their own pace, and made a response about the content of the subordinate clause of each experimental sentence.

Ambiguous: While the man/ coached / the woman/ attended/ the party by herself.

Unambiguous: While the man/ coached, / the woman/ attended/ the party by herself.

The garden path sentences contained an ambiguity at the verb in the subordinate clause (*coached*), with the initial interpretation involving taking the verb in the transitive form

and the subsequent ambiguous noun phrase as the direct object of the verb (*the man coached the woman*) (see Table 1.1 for a full set of examples). The control sentences unambiguously marked the subordinate verb in the intransitive form through presence of an accompanying comma. The ambiguous noun phrase was constructed to be a plausible object of the subordinate verb in all sentences. In the above example, the initial garden path interpretation has *the woman* – the entity denoted by the ambiguous noun phrase as the patient of *coached* – the action expressed in the subordinate clause. In successful semantic reinterpretation of the sentence, *the woman* is reassigned as the agent (i.e., doer) of *attended* – the action expressed in the main clause.

Besides ambiguity (presence vs. lack of an accompanying comma at the subordinate verb), the structural preference of the subordinate verb (i.e., verb bias) and the real world plausibility of the initial interpretation given the information subsequently presented in the main clause were manipulated. Verb bias encodes the relative frequency that alternative structures appear with a verb. The materials included *transitive-biased* verbs – verbs that appear frequently with a transitive structure and infrequently with an intransitive structure, and *equi-biased* verbs – verbs that appear as frequently with a transitive structure as with an intransitive structure. Verb bias was determined using the Galh, Jurafsky, and Roland's (2004) norm. Although it is controversial whether verb bias affects early stages of processing when the initial structural decision is made and pursued (see Adam, Clifton, & Mitchell, 1998; Garnsey, Pearlmutter, Myers, & Loocky, 1997, Trueswell, 1996; Trueswell & Kim, 1998; Trueswell, Tanenhaus, & Kello, 1993; Wilson & Garnsey, 2009 for positive evidence; see Ferreira & Henderson, 1990, 1991; Mitchell, 1987, Pickering, Traxler, & Crocker, 2000; Traxler, 2005; van Gompel & Pickering,

2001 for negative evidence), effects of verb bias on later stages of processing (e.g., the reinterpretation process) are more clear (e.g., Ferreira & Henderson, 1990, 1991; Frazier & Clifton, 1998; Traxler, 2005). When the verb is biased toward the initial interpretation, the demand on executive control may increase, as the verb's bias adds as a cue that could interfere with reinterpretation. Therefore, reinterpretation of object/subject garden path sentences should be harder with transitive-biased than with equi-biased verbs.

In the third and final manipulation of the materials, the semantic information subsequently presented in the main clause either matched or did not match with the initial object interpretation, making the initial interpretation *plausible* or *implausible*, given real world knowledge (see examples).

Implausible: While the man coached (,) the woman attended the party by herself.

Plausible: While the man coached (,) the woman attended to the helpful advice.

In the above example, it is not plausible that the woman was being coached because she was attending a party by herself. However, that the woman attended to some helpful advice is consistent with the idea that she was being coached. Whereas the mismatching semantic information serves as a semantic error cue against the initial interpretation, the matching semantic information adds as another cue in favor of the initial interpretation. The demand on executive control should accordingly be higher in the matching (plausible) condition than in the mismatching (implausible) condition, and semantic reinterpretation should be harder in the plausible than in the implausible condition.

To summarize, the ambiguity at the subordinate verb, the structural preference of the subordinate verb, and the plausibility of the initial interpretation given later semantic

information were varied in the materials. There were two other factors that were constant in the materials and worked in favor of the initial interpretation. First, the ambiguous noun phrase was a plausible object of the subordinate verb in all sentences. Second, a general preference for the transitive interpretation may exist due to overall frequency of the transitive structure in the language (e.g., Mitchell et al., 1995) or due to general parsing heuristics (e.g., late closure; Frazier & Fodor, 1978).

Measure of semantic recovery

Since ambiguous sentences lead to a misinterpretation while unambiguous control sentences presumably do not, reinterpretation should be specific to the ambiguous sentences. The extent to which reading time increases at the critical region (see below) and/or comprehension accuracy decreases in an ambiguous compared to a corresponding unambiguous condition reflects ability to reinterpret the meaning of garden path sentences. Ambiguity effects on comprehension accuracy (unambiguous accuracy minus ambiguous accuracy) and reading time at the critical region (ambiguous RT minus unambiguous RT) therefore served as measures of semantic recovery. Bigger ambiguity effects reflect poorer semantic recovery ability and vice versa.

Reading time measure. Based on Sturt (2007), it was assumed that semantic reinterpretation, if any, would mainly take place at the *final* region of the sentence where the semantic matching or mismatching cue was located. As explained above, between the two plausibility conditions, the plausible condition should place a greater demand on executive control. Between the two verb bias conditions, the transitive-biased verb condition should place a greater demand on executive control. Crossing between plausibility and verb bias, the demand on executive control during semantic

reinterpretation has the following rank ordering across the conditions (from lowest to highest): (1) implausible equi-biased verb, (2) implausible transitive-biased verb, (3) plausible equi-biased verb, and (4) plausible transitive-biased verb.

Four ambiguity effects on mean reading time for the final region of the sentence that corresponded to each of the four plausibility by verb bias conditions and an overall ambiguity effect collapsing across all conditions were calculated for each participant and individually correlated with each executive control RT measure across participants.

Accuracy measure. An overall effect of ambiguity on garden path interpretation accuracy and four specific ambiguity effects that corresponded to the plausibility by verb bias interaction were calculated for each participant and individually correlated with each executive control accuracy measure across participants.

Measure of verbal executive control

A measure of executive control for verbal materials was collected using the verbal Stroop task (Stroop, 1935). In the *incongruent* condition of the task, participants saw color words written in a mismatching ink color (the word RED written in blue ink), and must ignore the name of the word (red) and name aloud the color of the ink (blue). In the *neutral* condition, participants saw rows of asterisks written in color ink (***** written in blue ink) and named the color of the ink (blue). The demand on executive control is therefore high in the incongruent condition and minimal in the neutral condition. The extent to which naming latency increases and/or naming accuracy decreases in the incongruent compared to the neutral condition reflects verbal executive control ability. Bigger verbal Stroop effects on naming accuracy (neutral minus incongruent) and naming

latency (incongruent minus neutral) reflect poorer verbal executive control ability and vice versa.

Measure of nonverbal executive control

A measure of executive control for nonverbal materials was collected using a nonverbal version of the Stroop task (Hamilton & Martin, 2005). In the *incongruent* condition, participants saw pictures of arrows appearing at a location on the computer screen that mismatched the direction to which they were pointing (right-pointing arrow appearing on the left side of the screen); they must ignore the arrow's location (left) and respond based on its pointing direction (right). In the *neutral* condition, participants saw arrows appearing in the middle of the screen and responded based on its pointing direction. The demand on executive control is therefore high in the incongruent condition and minimal in the neutral condition. The extent to which response latency increases and/or response accuracy decreases in the incongruent compared to the neutral condition reflects nonverbal executive control ability. Bigger nonverbal Stroop effects on response accuracy (neutral minus incongruent) and response latency (incongruent minus neutral) reflect poorer nonverbal executive control ability and vice versa.

Predicted relationship between executive control and semantic recovery

According to the domain-specific executive control hypothesis, verbal executive control plays a role in semantic reinterpretation of garden path sentences, such that comprehenders who have better verbal executive control will be faster and/or more accurate in revising the initial semantic misinterpretation. This hypothesis therefore predicted a positive correlation between semantic recovery and verbal executive control

only. No correlation with nonverbal executive control would be found according to this hypothesis.

According to the domain-general executive control hypothesis, domain-general executive control plays a role in semantic reinterpretation of garden path sentences. This hypothesis would predict a positive correlation between semantic recovery and verbal executive control, *and* a positive correlation between semantic recovery and nonverbal executive control.

Alternatively, semantic reinterpretation of garden path sentences may be carried out through automatic processes. If that is the case, neither verbal nor nonverbal executive control should correlate with semantic recovery.

Methods

Participants

Fifty native speakers of English, 49 Rice University undergraduates and 1 graduate student from a nearby university in Houston, TX, participated in the experiment (half males half females, mean age = 19.5, $SD = 1.3$). Forty-seven of them participated in exchange for credit toward course requirements for undergraduate psychology classes at Rice University; the rest received monetary compensation at a rate of \$10 per hour. Two participants could not perform the verbal Stroop task due to color blindness and were dropped from all analyses.

Garden path sentence comprehension task

Materials. The experimental materials contained a manipulation of three factors – Plausibility, Verb Bias, and Ambiguity. Verb Bias was a between-item factor, as a verb could be classified as a member of one bias category only (either transitive-biased or

equi-biased). Plausibility and Ambiguity were fully crossed in the materials and were within-item factors. Examples of the materials are given in Table 1.1 (see Appendix A for a full set of experimental materials).

Table 1.1. Examples of materials for garden path sentence processing tasks in Experiment 1-3.

Equi-biased verb

Implausible, A: While the man/ coached / the woman/ attended/ the party by herself.

Implausible, U: While the man/ coached, / the woman/ attended/ the party by herself.

Plausible, A: While the man/ coached / the woman/ attended/ to the helpful advice.

Plausible, U: While the man/ coached, / the woman/ attended/ to the helpful advice.

Transitive-biased verb

Implausible, A: While the uncle/ visited / the child/ was/ missing him at home.

Implausible, U: While the uncle/ visited, / the child/ was/ missing him at home.

Plausible, A: While the uncle/ visited / the child/ was/ acting nice and quiet.

Plausible, U: While the uncle/ visited, / the child/ was/ acting nice and quiet.

Slashes indicate presentation region boundaries.

A: Ambiguous

Underlined regions were involved in reading time analysis.

U: Unambiguous

Forty experimental sentence items were constructed to contain a subordinate clause followed by a main clause. The noun phrase following the verb in the subordinate clause could be analyzed as the object of the subordinate verb or the subject of the main clause. Half of the items contained subordinate verbs that were equi-biased between the transitive (appearing with a direct object) and intransitive (appearing without a direct object) structure, the other half contained subordinate verbs that were biased toward the transitive structure. Verb bias was determined using the Gahl, Jurafsky, and Roland's (2004) norm. Verbs were classified as *equi-biased* if they occurred as often with a direct object as without one (difference not greater than 15%) (estimated mean frequency = 27% with a direct object and 27% without a direct object). Verbs were classified as *transitive-biased* if they occurred at least twice as often with as without a direct object (estimated mean frequency = 67% with a direct object and 12% without a direct object).

Four sentence versions were created for each item in the two bias conditions that fully crossed plausibility with ambiguity (plausible ambiguous, plausible unambiguous, implausible ambiguous, implausible unambiguous). The plausibility manipulation was localized at the final region of the sentence, which provided semantic information that was either consistent or inconsistent with the initial object interpretation, making the initial interpretation *plausible* or *implausible*. The ambiguity manipulation was localized at the subordinate verb, which was *ambiguous* if it appeared without an accompanying comma and *unambiguous* if it appeared with an accompanying comma to mark the intransitivity usage of the subordinate verb. The subsequent noun phrase was a plausible object of the subordinate verb in all experimental sentences (which all ultimately resolved against the object interpretation of the noun phrase).

Four lists of materials were created with 20 equi-biased verbs and 20 transitive-biased verbs included in each list. Each verb appeared in only one of the four plausibility by ambiguity conditions in a list but, across the four lists, each verb would have appeared once in each condition (the same number of items were included in each condition). Following each experimental sentence, a prompt appeared to ask participants about the information presented in the subordinate clause (e.g., *Tell me about the uncle*). Besides the experimental sentences, 120 filler sentences were included in each list (see Appendix B for examples). Forty filler sentences contained subordinate verbs that appeared without a direct object (half with and half without an accompanying comma, all with comprehension questions asking for information presented in the main clause). The remaining 80 filler sentences contained subordinate verbs that appeared with a direct object (half with comprehension questions asking about the subordinate clause, half about

the main clause). Participants were randomly assigned to only one of the four material lists.

Stimulus rating. To verify that the characteristics of the experimental sentences were as intended, new groups of Rice University undergraduates who were native speakers of English were recruited to rate the stimuli on (a) the acceptability of the subordinate verbs in the intransitive form, (b) the plausibility of the subsequent noun phrase as the direct object of the subordinate verb, and (c) the plausibility of the object interpretation of the ambiguous noun phrase given semantic information presented in the main clause. A summary of the properties of the materials is presented in Table 1.2.

Table 1.2. Properties of garden path sentence materials used in Experiment 1-3 and 5-6.

Verb bias	Mean frequency		(a) Acceptability	(b) Plausibility	(c) Plausibility DO given main clause	
	transitive	intransitive	Sub-V	DO	Plausible	Implausible
EQ	27%	27%	92%	6.54	5.83	1.99
TR	67%	12%	94%	6.62	5.85	2.02

EQ: equi-biased verb TR: transitive-biased verb

(a) Acceptability of the subordinate verbs in the intransitive form.

(b) Plausibility of the subsequent noun phrase as the direct object of the subordinate verb.

(c) Plausibility of the direct object interpretation of the noun phrase subsequent to the subordinate verb, given later semantic information in the main clause.

In the acceptability judgment task, participants saw only the subordinate clause of experimental sentences in which the subordinate verb was unambiguously marked in the intransitive form through presence of an accompanying comma (*While the uncle visited,...*). They indicated whether those and other ungrammatical filler fragments (e.g., *While the teenager damaged,...*) were acceptable. As intended, the fragments involving the equi-biased and transitive-biased verbs were judged to be acceptable by the majority of the participants and the mean rates of acceptability were similar across the equi-biased (92%) vs. transitive-biased verbs (94%) ($n = 22$, $t < 1$).

In the direct object plausibility rating task, participants saw simple sentences consisting of the subject of the subordinate clause, the subordinate verb, and the subsequent noun phrase of the experimental sentences (e.g., *The uncle visited the child*). They rated the plausibility of the noun phrase as the object of the verb in the experimental and other filler sentences (e.g., *The child begged the corpse*) on a scale from 1 – highly implausible to 7 – highly plausible. As intended, the plausibility rating for the experimental items was high and similar across the equi-biased verb (mean plausibility rating = 6.54) and transitive-biased verb items (mean plausibility rating = 6.62) ($n = 22$, $t < 1$).

Finally, in the plausibility rating of the object interpretation of the noun phrase subsequent to the subordinate verb, given matching vs. mismatching semantic information presented in the main clause, participants saw the entire experimental sentence but with a subject pronoun added to the main clause (e.g., *While the uncle visited the child, she was missing him at home*). They rated the plausibility of the subject pronoun (e.g., *she*) to refer to the same entity that was expressed in the preceding object noun phrase (e.g., *the child*) in the experimental and other filler sentences (e.g., *While the friend was blaming the partner, he listened without saying a word*) on a scale from 1 – highly implausible to 7 – highly plausible. As intended, the sentences containing matching semantic information in the main clause were judged to be more plausible (mean plausibility rating = 5.84) than those containing mismatching semantic information (mean plausibility rating = 2.00) ($n = 22$, $p < .001$). Further, the semantically matching sentences that contained equi-biased verbs were rated as plausible as those that contained transitive-biased verbs (mean plausibility rating = 5.83 vs. 5.85, $t < 1$), and the

semantically mismatching sentences that contained equi-biased verbs were rated as implausible as those that contained transitive-biased verbs (mean plausibility rating = 1.99 vs. 2.02, $t < 1$).

Procedure. The sentences were presented one region at a time in the middle of a computer screen. Participants read each region and pressed a button to proceed to the next region. The duration of time between the two button presses was recorded as the reading time for the region. Following the end of each sentence, a command asking for information presented in the sentence appeared on the screen and participants answered by typing in their responses. The order of the trials was pre-randomized and fixed for all participants.

Verbal Stroop Task

The materials for this task were provided by Akira Miyake (University of Colorado, Boulder). Participants saw stimuli presented on the computer screen and named aloud the color of each stimulus as quickly as possible. In the neutral condition (72 trials), they saw asterisks presented in blue, green, orange, purple, red, or yellow. In the congruent condition (12 trials), they saw color words presented in the same color ink. In the incongruent condition (60 trials), they saw color words presented in a different ink color. Naming latencies were recorded with a voice key. The order of the trials was pre-randomized and fixed for all participants.

Nonverbal Stroop Task

Participants saw pictures of arrows pointing to the left, right, up, or down and responded by pressing as quickly as possible the appropriate key that corresponded to the pointing direction of the arrow. In the neutral condition (36 trials), the arrows appeared in

the middle of the screen. In the congruent condition (36 trials), the arrows appeared at a location that matched with their pointing direction (left-pointing arrow appearing at the left side of the screen). In the incongruent condition (36 trials), the arrows appeared at a location that mismatched their pointing direction (e.g., left-pointing arrow appearing at the right side of the screen). The order of the trials was randomized for each participant.

Results

Data are reported for each of the tasks and for the correlations between the measures of interest.

Garden path sentence comprehension task

Comprehension accuracy. Prior research has shown that comprehenders are more likely to make the object interpretation (i.e., to make incorrect responses) when the sentence is ambiguous than when the sentence is unambiguous, and when the object interpretation is plausible than when it is implausible (e.g., Christianson et al., 2001). Given those findings, a main effect of Ambiguity and a main effect of Plausibility on comprehension accuracy were expected. Additionally, a reduction in accuracy due to ambiguity may be greater when the initial object interpretation is plausible than when it is implausible (it may be more difficult to revise initial interpretations that are plausible). If that is case, an interaction between Plausibility and Ambiguity would be found. Finally, ambiguity effects may be modulated by verb bias, such that reinterpretation is more difficult and less likely to succeed for transitive-biased verbs than for equi-biased verbs (leading to a larger ambiguity effect for transitive verbs). If that is case, interaction effect involving Verb Bias and Ambiguity should be found.

Responses to the experimental sentences were coded by the author into three categories: *verb* (e.g., the uncle visited or he visited), *verb plus object* (e.g., the uncle/he visited the child), and *other* (e.g., omitted responses, responses with information presented in the main clause instead of the targeted subordinate clause). From a total of 1920 responses, 86.0% of the responses were coded into the *verb* category, 9.9% in the *verb plus object* category, and 4.1% in the *other* category. Response accuracy was calculated by dividing the number of *verb* responses over the number of all responses in each condition for each participant and each item. The data were then entered into a 2 (Plausibility: plausible vs. implausible) x 2 (Verb Bias: equi-biased vs. transitive-biased) x 2 (Ambiguity: ambiguous vs. unambiguous) within-subject ANOVA in the subject analysis, and a 2 x 2 x 2 mixed ANOVA in the item analysis (Plausibility and Ambiguity were within-item factors and Verb Bias was a between-item factor). A summary of the means and standard errors for each condition and the *F*-statistics from the ANOVAs is presented in Table 1.3.

Table 1.3. Experiment 1 – garden path comprehension among younger comprehenders: Mean accuracy (and standard errors) by condition and *F*-statistics.

Condition	Accuracy	Variable	F _{1,47}	F _{2,38}
Implausible EQ ambiguous	88% (2%)	Plausibility	31.93**	28.79**
Implausible EQ unambiguous	90% (2%)	Verb bias	2.82	<1
Implausible TR ambiguous	94% (2%)	Ambiguity	3.76†	3.72†
Implausible TR unambiguous	95% (2%)	Plausibility x Verb bias	5.15*	2.07
		Plausibility x Ambiguity	1.22	1.4
Plausible EQ ambiguous	78% (4%)	Verb Bias x Ambiguity	<1	<1
Plausible EQ unambiguous	83% (3%)	Plausibility x VB x Am	<1	<1
Plausible TR ambiguous	78% (3%)			
Plausible TR unambiguous	82% (3%)			
EQ: equi-biased verb		† indicates $p < .10$	* $p < .05$	** $p < .01$
TR: transitive-biased verb		VB x Am: Verb bias x Ambiguity		

As expected, the participants were significantly less accurate when the initial interpretation was plausible (80% correct) than when it was implausible (92% correct)

(main effect of Plausibility: $F_1(1, 47) = 31.93, p < .001$, $F_2(1, 38) = 28.79, p < .001$).

Also consistent with the expectation, the participants tended to be less accurate when the sentence was ambiguous (85% correct) than when it was unambiguous (87% correct), though the main effect of Ambiguity was only marginal by subjects and items ($F_1(1, 47) = 3.76, p = .06$, $F_2(1, 38) = 3.72, p = .06$). The ambiguity effect was similar for plausible vs. implausible sentences (Plausibility x Ambiguity interaction: $F_s < 1.5$). Finally, verb bias did not have a significant influence on the final interpretation of the sentences (no effects involving verb bias were significant by subjects or items in the ANOVA results).

Garden path reading time. Only trials with a correct response (i.e., *verb* responses) were included in the analyses. Three participants were removed due to low comprehension accuracy (less than 60% correct). Data from the remaining participants were analyzed for the two error cue regions – the main verb region (where the syntactic error cue was located) and the final region (where the semantic error cue was located).

Because word length varied among the conditions, effects of length on reading times were first removed. Following Ferreira and Clifton (1986), a linear regression predicting reading time from region length that included all experimental and filler sentences (reading times shorter than 200 ms and longer than 10 s were removed) was derived for each participant. Residual reading times with length effects removed that were 2.5 *SD* below or above a participant's mean time on each critical region for each experimental condition were replaced with the 2.5 *SD* cutoff value (less than 1% of the data were affected). For the main verb region, trimmed residual reading times were entered into a 2 (Verb Bias: equi-biased vs. transitive-biased) x 2 (Ambiguity: ambiguous vs. unambiguous) within-subject ANOVA in the subject analysis and a 2 x 2 mixed

ANOVA in the item analysis. For the final region, trimmed residual reading times were entered into a 2 (Plausibility: plausible vs. implausible) x 2 (Verb Bias: equi-biased vs. transitive-biased) x 2 (Ambiguity: ambiguous vs. unambiguous) within-subject ANOVA in the subject analysis and a 2 x 2 x 2 mixed ANOVA in the item analysis.

The mean residual reading times and standard errors for each condition and the *F*-statistics from the ANOVAs for the main verb region are presented in Table 1.4. Based on prior research (e.g., Christianson et al., 2001), it was expected that the participants would initially adopt the incorrect object interpretation in the ambiguous, but not in the unambiguous, condition. As a result, increased processing difficulty should occur in the ambiguous condition when the syntactic error cue arrived at the main verb, leading to longer reading time in the ambiguous vs. unambiguous condition at this region. As indicated previously, there is mixed evidence regarding whether verb bias interacts with ambiguity resolution at an early stage of ambiguity processing (e.g., Garnsey et al., 1997; Ferreira & Henderson, 1990). If early, a significant interaction between verb bias and ambiguity should be found at the main verb region. If late, the interaction effect would be found at the final region of the sentence.

Table 1.4. Experiment 1 – garden path comprehension among younger comprehenders: Mean residual reading times (and standard errors) by condition at the main verb region and *F*-statistics.

Condition	RT	Variable	F _{1,44}	F _{2,38}
Equi-biased ambiguous	257 (47)	Verb Bias	<1	<1
Equi -biased unambiguous	-73 (29)	Ambiguity	59.49**	67.72**
Transitive-biased ambiguous	246 (37)	Verb Bias x Ambiguity	<1	<1
Transitive -biased unambiguous	-93 (23)			

** indicates $p < .01$

The results confirmed that the participants initially adopted the incorrect interpretation. Collapsing across verb bias conditions, they were approximately 300 ms slower on average to read the main verb of the ambiguous vs. unambiguous sentences

(significant main effect of Ambiguity: $F_1(1, 44) = 59.49, p < .001, F_2(1, 38) = 67.72, p < .001$). The ambiguity effect was slightly but not significantly larger on average for the transitive-biased verb items (339 ms effect, compared to 330ms effect for the equi-biased verb items; Verb Bias x Ambiguity interaction: $F_s < 1$). Finally, collapsing across ambiguity conditions, the participants read the main verb at a similar speed in the transitive-biased vs. equi-biased verb condition (main effect of Verb Bias: $F_s < 1$).

For the final region, the mean reading times and standard errors for each condition and the F -statistics from the ANOVAs are presented in Table 1.5. Based on Sturt (2007), the mismatching semantic cue in the implausible conditions was expected to trigger semantic reinterpretation in the ambiguous condition, leading to longer reading time in the ambiguous vs. unambiguous condition (in the unambiguous condition the correct interpretation had presumably been adopted and no reinterpretation was needed). Given that the later information was consistent with the initial interpretation in the plausible conditions, semantic reinterpretation may fail to be triggered in the ambiguous conditions in most comprehenders, leading to similar reading times on plausible sentences across the ambiguous vs. unambiguous conditions. Additionally, verb bias may or may not have an effect on reinterpretation. If it does (e.g., Ferreira & Henderson, 1991), Verb Bias and Ambiguity should significantly interact.

Table 1.5. Experiment 1 – garden path comprehension among younger comprehenders: Mean residual reading times (and standard errors) by condition at the final region and F -statistics.

Condition	Accuracy	Variable	$F_{1,44}$	$F_{2,38}$
Implausible EQ ambiguous	373 (74)	Plausibility	1.89	<1
Implausible EQ unambiguous	138 (66)	Verb bias	3.41†	<1
Implausible TR ambiguous	229 (80)	Ambiguity	<1	<1
Implausible TR unambiguous	389 (103)	Plausibility x Verb bias	4.97*	1.22
		Plausibility x Ambiguity	<1	<1
Plausible EQ ambiguous	263 (90)	Verb Bias x Ambiguity	1.89	3.40†
Plausible EQ unambiguous	370 (112)	Plausibility x Verb Bias x Am	4.23*	6.19*
Plausible TR ambiguous	105 (53)			
Plausible TR unambiguous	86 (82)			

EQ: equi-biased verb Am: Ambiguous † indicates $p < .10$ * indicates $p < .05$
 TR: transitive-biased verb

There was a significant three-way interaction between Plausibility, Verb Bias, and Ambiguity ($F_1(1, 44) = 4.23, p = .05, F_2(1, 38) = 6.19, p = .02$). To further assess the interaction, paired comparisons were done that contrasted ambiguous vs. unambiguous reading times across each of the four plausibility by verb bias conditions. Consistent with the expectation that the mismatching semantic cue would be effective in triggering semantic reinterpretation and the finding that verb bias affects reinterpretation (e.g., Ferreira & Henderson, 1991), reading time was significantly longer for the ambiguous items in the least interfering, implausible equi-biased verb condition (ambiguity effect size = 235 ms, $t_1(44) = 2.45, p = .02, t_2(19) = 2.67, p = .02$), suggesting that semantic reinterpretation was carried out by most of the comprehenders in this condition. None of the other paired comparisons were significant ($ts < 1.8$), suggesting that semantic reinterpretation failed to be engaged by most comprehenders in those conditions. None of the other effects were significant by both subjects and items in the ANOVA results.

Verbal Stroop

The participants were significantly less accurate in the incongruent (97% correct) than in the neutral condition (99% correct) ($t(47) = 4.96, p < .001$). Trials that were

incorrect and trials in which naming latency was shorter than 200 ms or longer than 4000 ms were removed in RT analysis (8 trials were excluded). Remaining naming latencies that were 2.5 *SD* below or above a participant's mean latency for each condition were replaced with the 2.5 *SD* cutoff value (2.5% of the data were affected). Results showed that the participants were significantly slower to name target colors in the incongruent (780 ms) than in the neutral condition (637 ms) ($t(47) = 18.68, p < .001$).

Nonverbal Stroop

The participants showed perfect accuracy in both the incongruent and neutral condition. Trials in which response latency was shorter than 200 ms and longer than 4000 ms were removed in RT analysis (no trials were excluded). Remaining correct response latencies that were 2.5 *SD* below or above a participant's mean latency for each condition were replaced with the 2.5 *SD* cutoff value (2.8% of the data were affected). Results showed that participants were significantly slower to make responses in the incongruent (576 ms) than in the neutral condition (511 ms) ($t(47) = 13.54, p < .001$).

Relationship between semantic recovery and executive control

The means, standard deviations, and reliability estimates for the RT and accuracy measures are respectively presented in Table 1.6a and 1.6b. Reliability estimates were calculated using the random split-half correlation method and adjusted by the Spearman-Brown prophecy formula. As shown in the Table 1.6a, the semantic recovery RT measures produced markedly lower reliability estimates than the other measures.

Table 1.6a. Experiment 1 – correlation across younger comprehenders: Means, standard deviations, and reliability estimates for RT measures (final region).

RT measure	Mean	SD	Reliability
Overall ambiguity effect	-3	433	0.12
Ambiguity effect - implausible equi-biased	235	642	0.17
Ambiguity effect - implausible transitive-biased	-160	900	0.30
Ambiguity effect - plausible equi-biased	-108	798	0.14
Ambiguity effect - plausible transitive-biased	20	698	0.37
Verbal Stroop effect	146	54	0.81
Nonverbal Stroop effect	63	32	0.50

Table 1.6b. Experiment 1 – correlation across younger comprehenders: Means, standard deviations, and reliability estimates for accuracy measures.

Accuracy measure	Mean	SD	Reliability
Overall ambiguity effect	3%	10%	0.06 ¹
Ambiguity effect - implausible equi-biased	2%	20%	0.62
Ambiguity effect - implausible transitive-biased	1%	16%	0.71
Ambiguity effect - plausible equi-biased	5%	19%	0.47
Ambiguity effect - plausible transitive -biased	4%	21%	0.64
Verbal Stroop effect	3%	4%	0.56

¹There were several (insignificant) negative correlations between the specific ambiguity effects, which could be responsible for the low reliability of the overall ambiguity effect.

For correlation analysis, observations that had a studentized t value greater than |3| were identified as outliers and excluded from analysis. In the correlations on accuracy data, two observations were identified as outliers in the correlation between verbal executive control and the overall semantic recovery measure. In the correlations on RT data, the same participant was identified as an outlier in four correlations: (i) the overall semantic recovery measure and the verbal executive control measure, (ii) the overall semantic recovery measure and the nonverbal executive control measure, (iii) the semantic recovery measure for plausible transitive-biased verb condition and the verbal executive control measure, and (iv) and the semantic recovery measure for implausible transitive-biased verb condition and the nonverbal executive control measure. The correlation results are summarized in Table 1.7.

Table 1.7. Experiment 1 – correlation across younger comprehenders: Relationship between executive control and semantic recovery.

Measure	Semantic recovery	Executive control	
		Verbal	Nonverbal
Accuracy	Overall ambiguity effect	0.12	
	Ambiguity effect - implausible equi-biased	0.03	
	Ambiguity effect - implausible transitive-biased	0.20	
	Ambiguity effect - plausible equi-biased	-0.04	
	Ambiguity effect - plausible transitive-biased	0.18	
RT	Overall ambiguity effect	0.19	-0.07
	Ambiguity effect - implausible equi-biased	0.13	0.22
	Ambiguity effect - implausible transitive-biased	0.26†	-0.09
	Ambiguity effect - plausible equi-biased	0.36*	-0.14
	Ambiguity effect - plausible transitive-biased	-0.32*	-0.11

† indicates $p < .10$.

* indicates $p < .05$.

Significant correlations were found in the two conditions involving plausible garden path sentences that were hypothesized to place the highest demands on executive control. First, there was a significant positive correlation between verbal executive control and semantic recovery in the plausible equi-biased verb condition ($r = .36$, $p = .01$) (see Figure 1a). Second, in the most demanding, plausible transitive-biased verb condition, there was a significant *negative* correlation between verbal executive control and semantic recovery ($r = -.32$, $p = .04$) (see Figure 1b).

A pattern of a positive relationship between verbal executive control and garden path recovery was also found for the two less demanding, implausible garden path conditions. First, in the least demanding, implausible equi-biased verb condition, the correlation was positive ($r = .19$) and not significant ($p = .41$). Second, in the more demanding, implausible transitive-biased verb condition, the correlation was positive ($r = .26$), and marginally significant ($p = .08$).

No significant correlations were found between the nonverbal executive control measure and any of the semantic recovery measures ($ps > .15$) and no significant correlations were found for the accuracy measures ($ps > .18$).

Figure 1a. Experiment 1: Relationship between verbal Stroop effect and ambiguity effect on reading time in the plausible, equi-biased verb condition across participants.

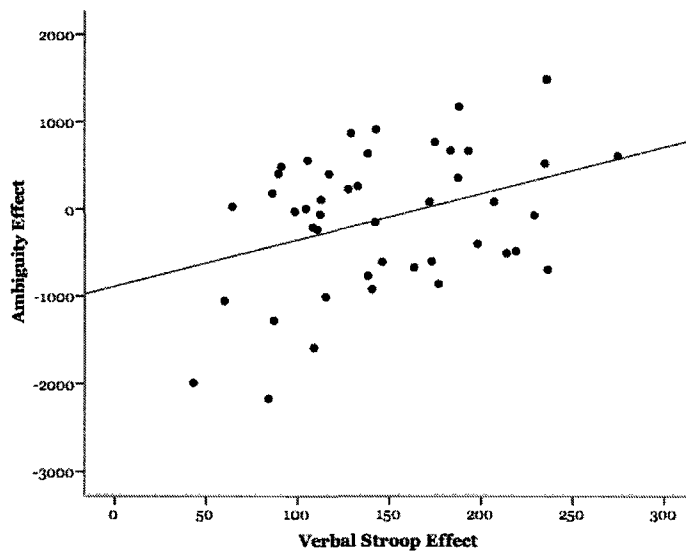
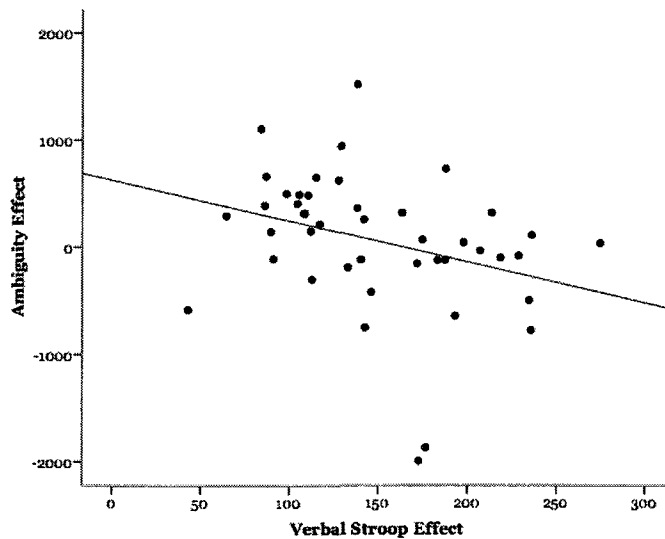


Figure 1b. Experiment 1: Relationship between verbal Stroop effect and ambiguity effect on reading time in the plausible, transitive-biased verb condition across participants.



Discussion of Experiment 1

The results of this experiment are generally consistent with the domain-specific executive control hypothesis of garden path reinterpretation. The semantic recovery RT measures correlated significantly with the verbal and not with the nonverbal executive control measure. The significant correlations were specifically found in the garden path comprehension conditions that were hypothesized to place higher demands on executive control. The positive correlation in the plausible equi-biased verb condition shows that comprehenders with poorer verbal executive control indeed need to take more time to revise the meaning of garden path sentences.

Unexpectedly, though not necessarily inconsistent with the executive control hypothesis, the semantic recovery measure significantly and negatively correlated with the verbal executive control measure in the most demanding, plausible transitive-biased verb condition. A plausible interpretation of this result is that, instead of taking more time to recover from the initial misinterpretation, participants with poorer verbal executive control tend to not engage in recovery or terminate the recovery process sooner than those with better verbal executive control when most cues in the sentence (e.g., general structural preference, the verb's preference, plausibility of the ambiguous noun phrase as the object of the verb, later matching semantic information) support the initial garden path misinterpretation. An inspection of the data reveals that all but four participants showed a positive ambiguity effect for the transitive-biased verb sentences at the *main verb* region, suggesting that those with poorer verbal executive control also engaged in some reanalysis effort (presumably mainly at the syntactic level). However, when the initial interpretation received converging sources of support from both verb bias and later

matching semantic information, those with poorer executive control simply did not spend the time necessary to revise the initial semantic misinterpretation.

Since the reading time data had been filtered for incorrect responses, questions arise as to why the comprehenders with poorer verbal executive control spent less time than did the comprehenders with better verbal executive control on semantic reinterpretation and, yet, they, too, managed to produce the correct responses. One possible explanation for this apparently contradictory aspect of the data is that the responses the participants made in this task were not completely semantically based (thus, accuracy rates did not totally reflect success in semantic revision). Given that almost all of the participants were able to guess, as suggested by their sensitivity to the syntactic disambiguating cue at the main verb, that the clausal boundary was supposed to be located at the subordinate verb, they might have conservatively generated comprehension responses based on the boundary. Consistent with this explanation is the high accuracy rate found in this study (approximately 78% correct for the plausible ambiguous conditions vs. approximately 30% correct rate found in a recent study that used a paraphrasing procedure to tap garden path comprehension (Patson et al., 2009)). If the responses had been solely semantically based, the accuracy rate in this study should have been substantially lower for the plausible conditions (closer to the 30% accuracy rate found in Patson et al. (2009)).

Contrary to some previous studies (e.g., Garnsey et al., 1997) but consistent with some other studies (e.g., Ferreira & Henderson, 1990; Traxler, 2005), the interaction between verb bias and ambiguity were found at a later stage of ambiguity resolution (at the final region of the sentence) in this experiment. That interaction effect and the

opposite direction of the correlations involving equi- vs. transitive-biased verbs are consistent with the claim that verb bias has a powerful effect on the garden path reinterpretation process (e.g., Ferreira & Henderson, 1990).

Experiment 2. Garden Path Sentence Comprehension: A Follow-up

Experiment 2 used a different procedure than that used in Experiment 1 to examine comprehension of garden path sentences. In this experiment, participants read object/subject garden path and unambiguous control sentences (e.g., *While the uncle visited (,) the child was missing him at home*) and responded to yes/no comprehension questions about the object interpretation of the ambiguous noun phrase (e.g., *Did the uncle visit the child?*). The materials contained the same Plausibility x Verb Bias x Ambiguity manipulations as in Experiment 1.

This experiment was done to address the concern regarding the higher than expected accuracy rate for the conditions in which later semantic information matched with the initial interpretation (the plausible conditions). The match should have encouraged maintenance of the initial interpretation, thus producing incorrect responses on a large portion (perhaps at least half or more) of the trials. However, on average the participants made incorrect responses in only about 20% of the trials in Experiment 1. As discussed previously, the testing procedure in that experiment might have been ineffective in probing the participant's final semantic interpretation of the sentences. The more direct yes/no questioning method was used in this experiment to address that concern. If the conjecture is correct, the accuracy rate for the plausible conditions should be substantially lower in this experiment than in Experiment 1.

Participants

Twenty native speakers of English who were Rice University undergraduates performed this task as part of a psycholinguistic experiment that included other unrelated tasks (12 females, 8 males, mean age = 18.8, $SD = 1.2$). They participated in exchange for course credits toward undergraduate psychology classes at Rice University.

Materials

The experimental materials in this experiment were the same as those used in Experiment 1. As in Experiment 1, the materials were varied in three factors – Plausibility, Verb Bias, and Ambiguity. Plausibility and Ambiguity were within-item factors while Verb Bias was a between-item factor. In this experiment, each sentence was followed by a yes/no comprehension question that directly asked about the object interpretation of the noun phrase subsequent to the subordinate verb (see examples below). As in Experiment 1, 120 filler sentences were also included (half requiring a “yes” response half “no”).

Equi-biased verb

Implausible, A: While the man/ coached / the woman/ attended/ the party by herself.

Implausible, U: While the man/ coached, / the woman/ attended/ the party by herself.

Plausible, A: While the man/ coached / the woman/ attended/ to the helpful advice.

Plausible, U: While the man/ coached, / the woman/ attended/ to the helpful advice.

Question: *Did the man coach the woman?*

Transitive-biased verb

Implausible, A: While the uncle/ visited / the child/ was/ missing him at home.

Implausible, U: While the uncle/ visited, / the child/ was/ missing him at home.

Plausible, A: While the uncle/ visited / the child/ was/ acting nice and quiet.

Plausible, U: While the uncle/ visited, / the child/ was/ acting nice and quiet.

Question: *Did the uncle visit the child?*

Slashes indicate presentation region boundaries.

A: Ambiguous

Underlined regions were involved in reading time analysis.

U: Unambiguous

Procedure

The sentences were presented one region at a time in the middle of the computer screen. Participants read each region and pressed a button to proceed to the next region. The duration of time between the two button presses was recorded as the reading time for the region. Following the end of each sentence, a yes/no comprehension question appeared on the screen and participants answered by pressing the “yes” or “no” button. The order of the trials was pre-randomized and fixed for all participants.

Results

Responses to the experimental sentences were counted as correct if the participants rejected the object interpretation in their responses (a “no” response). As expected, the overall accuracy in this experiment was lower than in Experiment 1, averaging at 70% correct (vs. 86% correct in Experiment 1). The rate was substantially lower in the plausible (51% correct) than in the implausible condition (88% correct). Because an examination of reading time was desirable to gain information about on-line processing of the sentences but the data for the plausible trials would be severely reduced once incorrect trials were removed, *all* reading time data (correct and incorrect) were included in the reading time analysis for this experiment (the mean untrimmed reading times in each condition for all trials and for correct vs. incorrect trials are presented in Table 2.4).

Comprehension accuracy. The accuracy data were examined in a 2 (Plausibility: plausible vs. implausible) x 2 (Verb Bias: equi-biased vs. transitive-biased) x 2 (Ambiguity: ambiguous vs. unambiguous) within-subject ANOVA in the subject analysis, and a 2 x 2 x 2 mixed ANOVA in the item analysis (Plausibility and Ambiguity

were within-item factors, Verb Bias was a between-item factor). The means and standard errors for each condition and the F -statistics from the ANOVAs are presented in Table 2.1.

Table 2.1. Experiment 2 – garden path comprehension among younger comprehenders: Mean accuracy (and standard errors) by condition and F -statistics.

Condition	Accuracy	Variable	$F_{1,19}$	$F_{2,38}$
Implausible EQ ambiguous	85% (5%)	Plausibility	106.47**	81.29**
Implausible EQ unambiguous	87% (5%)	Verb bias	1.58	<1
Implausible TR ambiguous	88% (4%)	Ambiguity	<1	<1
Implausible TR unambiguous	91% (3%)	Plausibility x Verb bias	<1	<1
		Plausibility x Ambiguity	<1	<1
Plausible EQ ambiguous	51% (7%)	Verb Bias x Ambiguity	<1	<1
Plausible EQ unambiguous	48% (6%)	Plausibility x VB x A	<1	<1
Plausible TR ambiguous	53% (8%)			
Plausible TR unambiguous	53% (5%)			
EQ: equi-biased verb		VB x A: Verb Bias x Ambiguity	** indicates $p < .01$	
TR: transitive-biased verb				

The main effect of Plausibility, with lower accuracy in the plausible (51% correct) vs. implausible condition (88% correct), was significant (effect size = 37%; $F_1(1, 19) = 106.47, p < .001, F_2(1, 38) = 81.29, p < .001$). In contrast to Experiment 1, in which there was a marginally significant main effect of ambiguity, the participants in this experiment were as accurate in the ambiguous (69% correct) as in the unambiguous conditions (70% correct; main effect of Ambiguity: $F_s < 1$). No other effects were significant in the ANOVA results ($F_s < 1.58$).

Reading time. As in Experiment 1, to correct for variation in word length across conditions, a linear regression predicting reading time from region length that included all experimental and filler sentences (reading times shorter than 200 ms and longer than 10 s were removed) was derived for each participant (Ferreira & Clifton, 1986). Residual reading times with length effects removed that were 2.5 SD below or above a participant's mean time in each critical region for each experimental condition were

replaced with the 2.5 *SD* cutoff value (less than 3% of the data were affected). For the main verb region, trimmed residual reading times were entered into a 2 (Verb Bias: equi-biased vs. transitive-biased) x 2 (Ambiguity: ambiguous vs. unambiguous) within-subject ANOVA in the subject analysis and a 2 x 2 mixed ANOVA in the item analysis. For the final region, trimmed residual reading times were entered into a 2 (Plausibility: plausible vs. implausible) x 2 (Verb Bias: equi-biased vs. transitive-biased) x 2 (Ambiguity: ambiguous vs. unambiguous) within-subject ANOVA in the subject analysis and a 2 x 2 x 2 mixed ANOVA in the item analysis. The mean residual reading times and standard errors for each condition and the *F*-statistics from the ANOVAs for the main verb and the final region are presented in Table 2.2 and Table 2.3, respectively.

Table 2.2. Experiment 2 – garden path comprehension among younger comprehenders: Mean residual reading times (and standard errors) by condition at the main verb region and *F*-statistics.

Condition	RT	Variable	$F_{1,19}$	$F_{2,38}$
Equi-biased ambiguous	222 (46)	Verb Bias	14.66**	4.92*
Equi-biased unambiguous	-88 (45)	Ambiguity	33.36**	72.47**
Transitive-biased ambiguous	471 (64)	Verb Bias x Ambiguity	7.63*	4.66*
Transitive-biased unambiguous	-51 (45)			

* indicates $p < .05$

** indicates $p < .01$

As in Experiment 1, the reading time results showed that the participants had pursued the incorrect interpretation and experienced processing difficulty when the syntactic error cue arrived at the main verb. They were 416 ms slower to read the main verb in the ambiguous vs. unambiguous condition (main effect of Ambiguity: $F_1(1, 19) = 33.36, p < .001, F_2(1, 38) = 72.47, p < .001$). However, unlike in Experiment 1, there was a main effect of verb bias and an interaction between ambiguity and verb bias. Participants took longer to read the transitive-biased verb than to read the equi-biased verb items (143 ms difference in residual RTs) (significant main effect of Verb Bias: $F_1(1, 19) = 14.66, p = .001, F_2(1, 38) = 4.92, p = .03$). With regard to the interaction, the

effect of ambiguity was significantly larger for the items containing transitive-biased verbs (522 ms effect) than for those containing equi-biased verbs (311 ms effect) (significant Verb Bias x Ambiguity interaction: $F_1(1, 19) = 7.63, p = .01, F_2(1, 38) = 4.66, p = .04$). The greater ambiguity effect for the transitive verbs occurs earlier in this experiment (at the main verb here vs. at the final region in Experiment 1) and suggests that the participants experienced greater difficulty in revising the initial misinterpretation when the ambiguous verb strongly engendered that interpretation.

Table 2.3. Experiment 2 – garden path comprehension among younger comprehenders: Mean residual reading times (and standard errors) by condition at the final region and F -statistics.

Condition	RT	Variable	$F_{1,19}$	$F_{2,38}$
Implausible EQ ambiguous	298 (91)	Plausibility	<1	<1
Implausible EQ unambiguous	219 (74)	Verb Bias	1.52	<1
Implausible TR ambiguous	235 (91)	Ambiguity	3.21†	4.03†
Implausible TR unambiguous	243 (66)	Plausibility x Verb Bias	<1	<1
		Plausibility x Ambiguity	2.75	1.41
Plausible EQ ambiguous	428 (87)	Verb Bias x Ambiguity	3.79†	2.04
Plausible EQ unambiguous	138 (86)	Plausibility x VB x A	<1	<1
Plausible TR ambiguous	207 (72)			
Plausible TR unambiguous	137 (77)			

EQ: equi-biased verb

VB x A: Verb bias x Ambiguity

† indicates $p < .10$

TR: transitive-biased verb

At the final region, where the semantic error cue was located, the participants were 108 ms slower on average to in the ambiguous vs. unambiguous condition. The main effect of Ambiguity was marginal by subjects and by items ($F_1(1, 19) = 3.21, p = .09, F_2(1, 38) = 4.03, p = .05$). There was an overall pattern of a larger ambiguity effect in the less interfering equi-biased verb (185 ms effect) compared to the transitive-biased verb condition (31 ms effect). However, the interaction between Verb Bias and Ambiguity was only marginally significant by subjects ($F_1(1, 19) = 3.79, p = .07$) and not significant by items ($F_2(1, 19) = 2.04, p = .16$). None of the other effects were significant in the ANOVA results.

Table 2.4. Experiment 2 – garden path comprehension among younger comprehenders: Mean untrimmed reading times by condition for all trials and for correct vs. incorrect trials.

Critical region	Plausibility	Trial type	EQ	EQ	TR	TR
			ambiguous	unambiguous	ambiguous	unambiguous
Main verb	Both	All trials	976	666	1218	695
	Both	Correct trials	992	649	1308	717
	Both	Erroneous trials	941	702	1003	639
Final	Implausible	All trials	1637	1567	1470	1476
	Implausible	Correct trials	1567	1591	1415	1473
	Implausible	Erroneous trials	2035	1407	1869	1501
	Plausible	All trials	1819	1524	1483	1408
	Plausible	Correct trials	1885	1660	1588	1503
	Plausible	Erroneous trials	1751	1398	1366	1302

EQ: equi-biased verb

TR: transitive-biased verb

Discussion of Experiment 2

As expected, with the more direct testing procedure, the participants performed substantially less accurately on the plausible sentences in this experiment (51% correct in this experiment vs. 80% correct in Experiment 1; the results for the implausible conditions were comparable across the two experiments: 88% correct in this experiment vs. 92% correct in Experiment 1). These results are consistent with the speculation that the responses made in Experiment 1 were not driven by strictly semantic considerations. As mentioned previously, clause boundary may have been an important factor in driving the responses in Experiment 1.

Two other aspects of the current data are worth noting. First, the significant main effect of plausibility on comprehension accuracy suggests that under certain condition, such as when the noun phrase following the ambiguous verb is a plausible object of the verb and later semantic information does not contradict the object interpretation of the noun phrase, presence of the comma cue is not totally effective in discouraging comprehenders from making the object interpretation of the noun phrase. Second, there

was clear evidence in this experiment that verb bias modulated the extent of the ambiguity effect at the syntactic error cue (main verb region). Further study is needed to specify the conditions that give rise to an early (found in this experiment) vs. late (found in Experiment 1) emergence of the verb bias modulation effect.

Experiment 3. Executive Control and Syntactic Recovery

Experiment 3 focused on syntactic reinterpretation of garden path sentences. The experimental materials and the procedure were the same as those in Experiment 1, with the exception that participants in this experiment made grammaticality judgments about the sentences. Responses were scored as correct if participants judged the experimental sentences to be grammatical and incorrect if they judged them to be ungrammatical. For the ambiguous sentences (e.g., *While the man coached the woman attended to the helpful advice*), participants were expected to initially take the subordinate verb (*coached*) to be in the transitive form and take the subsequent noun phrase (*the woman*) as the object of the verb. Following Sturt (2007), syntactic revision was assumed to occur upon arrival of the syntactic error cue at the main verb (*attended*) (reinterpreting the subordinate verb to be in the intransitive form and the subsequent noun phrase as the subject of the main verb). If the participant fails to revise the initial syntactic interpretation of the garden path sentence, the main verb would still lack a grammatically licensed subject, thus causing him/her to perceive the sentence to be ungrammatical and consequently making a “no” response in the task. For the unambiguous sentences, participants were assumed to be able to avoid the initial object misinterpretation thanks to presence of the comma marker, thus taking the ambiguous noun phrase as the subject in the main clause and attaching the main verb to it without difficulty, and consequently judge the sentence to be grammatical.

A drop in the accuracy rate in the ambiguous compared to the unambiguous condition, therefore, reflects failure to revise the initial structural interpretation of the garden path sentence. As a preview of the results, the overall accuracy rate was lower in this experiment than in Experiment 1, averaging at approximately 63% correct. As in Experiment 2, reading time analysis was done on *all* (correct and incorrect) trials.

Measure of syntactic recovery

Recent evidence suggests that meaning could drive structural interpretation under certain circumstances (e.g., Ferreira, 2003; Kim & Osterhout, 2005). For example, given a sentence that starts with an unambiguous fragment like *The hearty meal was devouring...* (Kim & Osterhout, 2005), syntax unambiguously signals that *the meal* is the subject, thus doer, of *devouring*, giving the impression that the meaning is anomalous. However, semantic knowledge would suggest that *the meal* should be devoured, giving the impression that the structure is anomalous. Comprehenders have been found to interpret strings like the above according to semantic knowledge and react as if the structure was wrong (as indexed by the P600 effect on the comprehender's event-related brain potentials recorded while they were reading the sentences) (Kim & Osterhout, 2005).

If grammaticality judgments are based solely on grammatical considerations, then failed syntactic revision should lead to an incorrect response (the sentence is not grammatical) while successful reparsing of the garden path sentence should lead to a correct response (the sentence is grammatical). However, if aspects of meaning influence the perception of grammaticality, we may expect the sentences with mismatching semantic representations (e.g., later semantic information contradicts the initial

interpretation as in “While the uncle visited the child was missing him at home”) to be less likely perceived as grammatical than those containing matching semantic representations (e.g., “While the uncle visited the child was acting nice and quiet”). Note that this should be true only when *the comprehender still maintains the object interpretation* of the ambiguous noun phrase. If s/he correctly reinterprets the noun phrase as the subject of the main clause and the subordinate verb as being in the intransitive form (e.g., the uncle was visiting someone unspecified), then no mismatch with later semantic information should be present (e.g., while the uncle was visiting someone/somewhere unspecified, the child was missing him at home). In short, in the semantically mismatching conditions, reliance on shallow, matching vs. mismatching, semantic representations may lead to incorrect responses in the grammaticality judgment task (the sentence is ungrammatical).

However, in the condition in which later semantic information does not contradict the initial interpretation, reliance on semantics while making grammaticality judgments may lead to an inflated correct response rate. Assume that the comprehender inappropriately integrates later semantic representations with the initial object interpretation (e.g., while the uncle was visiting the child, she was acting nice and quiet), the sentence sounds fine semantically. If semantics influences the comprehender’s judgment of grammaticality (“the meaning of this sentence sounds fine, therefore the sentence must be grammaticality acceptable”), s/he would accept the sentence as grammatical but *for the wrong reason*. Coincidentally, successful revision of the syntactic structure of the garden path sentence would also lead to the judgment that the

sentence is grammatical. In the semantically matching conditions, then, a correct response may not be indicative of successful syntactic recovery.

In summary, the accuracy rate in this experiment may be *higher* for the *plausible* than for the implausible sentences, due to a conflation between semantics and syntax in producing correct responses in the plausible but not in the implausible condition. For the implausible condition, reliance on shallow semantics would lead to incorrect responses while successful syntactic revision would lead to correct responses. As in Experiment 1, an overall effect of ambiguity and four specific ambiguity effects corresponding to the plausibility by verb bias interaction on grammaticality judgment accuracy were calculated for each participant and separately correlated with the verbal and nonverbal executive control accuracy measure.

For the RT measure, an overall ambiguity effect and two specific ambiguity effects, for the transitive-biased vs. equi-biased verb condition, at the main verb region where the syntactic error cue was located (and where syntactic revision presumably occurred) were calculated for each participant and separately correlated with the verbal and nonverbal executive control RT measure.

Predicted relationship between executive control and syntactic recovery

According to the domain-specific executive control hypothesis, verbal executive control plays a role in the revision of the syntactic structure of garden path sentences. Consequently, comprehenders who have better verbal executive control would be more likely to succeed and/or faster in syntactic revision than those with poorer verbal executive control. This hypothesis therefore predicted a positive correlation between

syntactic recovery and verbal executive control only. No correlation with nonverbal executive control would be found according to this hypothesis.

According to the domain-general executive control hypothesis, domain-general executive control plays a role in syntactic recovery. This hypothesis predicted a positive correlation between syntactic recovery and verbal executive control, and a positive correlation between syntactic recovery and nonverbal executive control.

Alternatively, syntactic reinterpretation may be carried out through automatic processes. If that is the case, neither the verbal nor the nonverbal executive control would correlate with syntactic recovery.

Methods

Participants

Forty-eight native speakers of English who were Rice University undergraduates participated in the experiment (23 males, mean age = 19.9, $SD = 2.8$) in exchange for credit toward course requirements for undergraduate psychology classes at Rice University. Data from one participant were missing for the nonverbal Stroop task, leaving 47 participants available for data analysis of that task.

Tasks

Garden path grammaticality judgment. The experimental materials in this experiment were the same as those in Experiment 1. However, the filler materials were different for this experiment (see Appendix C for examples). Eighty filler sentences that had a similar sentence structure to the experimental sentences, half grammatical half ungrammatical, were included. All of the grammatical filler sentences had the subordinate verb appearing in the transitive form. Half of the ungrammatical filler

sentences were missing a subject in the subordinate clause (10 sentences) or in the main clause (10 sentences), the other half were missing a required object in the subordinate clause (10 sentences) or in the main clause (10 sentences). A comma was inserted at clause boundary for half of the fillers.

Participants read the sentences one phrase at a time at their own pace. Following the end of each sentence, a prompt appeared that asked them to judge the grammaticality of the sentence. Participants pressed one of two buttons to indicate that the sentence was grammatical (yes) or ungrammatical (no).

Verbal Stroop. The materials and the procedure of this task were the same as those in Experiment 1.

Nonverbal Stroop. The materials and the procedure of this task were the same as those in Experiment 1.

Results

For the garden path grammaticality judgment task, a “yes” response indicated that the participant accepted the sentence as grammatical and was scored as correct. Overall, the participants were 63% correct on the experimental trials and 87% correct on the filler trials. As in Experiment 2, *all* reading times (correct and incorrect) were included in the analysis of reading time (see Table 3.4 for mean untrimmed reading times in each condition for all trials and for correct vs. incorrect trials).

Garden path grammaticality judgment task

Grammaticality judgment accuracy. The mean accuracy and standard errors for each condition and the F -statistics on grammaticality judgment accuracy is presented in Table 3.1.

Table 3.1. Experiment 3 – garden path grammaticality judgment among younger comprehenders: Mean accuracy (and standard errors) by condition and F -statistics.

Condition	Accuracy	Variable	$F_{1,47}$	$F_{2,38}$
Implausible EQ ambiguous	53% (4%)	Plausibility	13.04**	4.6*
Implausible EQ unambiguous	68% (4%)	Verb bias	<1	<1
Implausible TR ambiguous	48% (5%)	Ambiguity	40.96**	83.03**
Implausible TR unambiguous	68% (4%)	Plausibility x Verb bias	4.39*	<1
		Plausibility x Ambiguity	<1	<1
Plausible EQ ambiguous	53% (4%)	Verb Bias x Ambiguity	<1	<1
Plausible EQ unambiguous	77% (4%)	Plausibility x Verb Bias x Am	1.32	2.05
Plausible TR ambiguous	61% (4%)			
Plausible TR unambiguous	78% (3%)			

EQ: equi-biased verb TR: transitive-biased verb * indicates $p < .05$ ** indicates $p < .01$

As expected, collapsing across plausibility and ambiguity, the participants were significantly less accurate in the ambiguous (54% correct) than in the unambiguous condition (73% correct) (main effect of Ambiguity: $F_1(1, 47) = 40.96, p < .001, F_2(1, 38) = 83.03, p < .001$). Consistent with the prediction regarding the influence of plausibility on grammaticality judgments, participants were more accurate in the plausible (67% correct) than in the implausible condition (59% correct) (significant main effect of Plausibility: $F_1(1, 47) = 13.04, p = .001, F_2(1, 38) = 4.60, p = .04$). No other effects were significant by both subjects and items in the ANOVA results.

Reading time. As in the previous experiments, to correct for variation in word length across conditions, a linear regression predicting reading time from region length that included all experimental and filler sentences (reading times shorter than 200 ms and longer than 10 s were removed) was derived for each participant (Ferreira & Clifton, 1986). Residual reading times with length effects removed that were 2.5 SD below or

above a participant's mean time on each critical region for each experimental condition were replaced with the 2.5 *SD* cutoff value (less than 3% of the data were affected). For the main verb region, trimmed residual reading times were entered into a 2 (Verb Bias: equi-biased vs. transitive-biased) x 2 (Ambiguity: ambiguous vs. unambiguous) within-subject ANOVA in the subject analysis and a 2 x 2 mixed ANOVA in the item analysis. For the final region, trimmed residual reading times were entered into a 2 (Plausibility: plausible vs. implausible) x 2 (Verb Bias: equi-biased vs. transitive-biased) x 2 (Ambiguity: ambiguous vs. unambiguous) within-subject ANOVA in the subject analysis and a 2 x 2 x 2 mixed ANOVA in the item analysis. The mean residual reading times and standard errors for each condition and the *F*-statistics from the ANOVAs for the main verb and the final region are presented in Table 3.2 and Table 3.3, respectively.

Table 3.2. Experiment 3 – garden path grammaticality judgment among younger comprehenders: Mean residual reading times (and standard errors) by condition at the main verb region and *F*-statistics.

Condition	RT	Variable	$F_{1,47}$	$F_{2,38}$
Equi-biased ambiguous	488 (62)	Verb Bias	1.09	<1
Equi-biased unambiguous	26 (25)	Ambiguity	60.98**	129.41**
Transitive-biased ambiguous	529 (58)	Verb Bias x Ambiguity	<1	<1
Transitive-biased unambiguous	31 (20)			

** indicates $p < .01$

Results showed that the participants were garden-pathed and needed to revise their syntactic interpretation when the syntactic error cue occurred at the main verb. They were significantly slower to read the main verb region in the ambiguous than in the unambiguous condition (overall ambiguity effect size = 480 ms; main effect of Ambiguity: $F_1(1, 47) = 60.98, p < .001, F_2(1, 38) = 129.41, p < .001$). The ambiguity effect was larger for the transitive-biased verb (498 ms effect) than for the equi-biased verb items (463 ms effect). However, the interaction between Verb Bias and Ambiguity

was not significant in this experiment ($F_s < 1$). Overall, the participants read the sentences containing the two verb bias types at a similar speed ($F_s < 1.1$).

Table 3.3. Experiment 3 – garden path grammaticality judgment among younger comprehenders: Mean residual reading times (and standard errors) by condition at the final region and F -statistics.

Condition	RT	Variable	$F_{1,47}$	$F_{2,38}$
Implausible EQ ambiguous	1189 (163)	Plausibility	6.71*	3.12†
Implausible EQ unambiguous	998 (159)	Verb Bias	1.80	<1
Implausible TR ambiguous	857 (133)	Ambiguity	<1	<1
Implausible TR unambiguous	762 (121)	Plausibility x Verb Bias	2.48	1.59
		Plausibility x Ambiguity	<1	<1
Plausible EQ ambiguous	731 (162)	Verb Bias x Ambiguity	<1	<1
Plausible EQ unambiguous	661 (168)	Plausibility x VB x A	<1	<1
Plausible TR ambiguous	719 (112)			
Plausible TR unambiguous	784 (162)			
EQ: equi-biased verb		† indicates $p < .10$ * indicates $p < .05$		
TR: transitive-biased verb		VB x A: Verb bias x Ambiguity		

The only effect found at the final region involved an overall effect of plausibility. Collapsing across ambiguity and verb bias conditions, participants were slower to read the final region of the sentences containing semantic information mismatching the object interpretation of the ambiguous noun phrase than to read the final region of the sentences containing matching semantic information (overall plausibility effect size = 228 ms; the main effect of Plausibility was significant by subjects, $F_1(1, 47) = 6.71$, $p = .01$, and marginally significant by items, $F_2(1, 38) = 3.12$, $p = .09$). Overall, the participants were slightly but not significantly slower in the ambiguous than in the unambiguous condition (overall ambiguity effect size = 72 ms, $F_s < 1$), suggesting that syntactic revision attempts did not spill over to this final region. No other effects were significant in the ANOVA results ($F_s < 2.5$).

Table 3.4. Experiment 3 – garden path grammaticality judgment among younger comprehenders: Mean untrimmed reading times by condition for all trials and for correct vs. incorrect trials.

Region	Plausibility	Trial type	EQ A	EQ U	TR A	TR U
Main verb	Both	All trials	1158	699	1172	708
	Both	Correct trials	1100	665	1197	726
	Both	Incorrect trials	1222	787	1142	658
Final	Implausible	All trials	3106	2901	2510	2640
	Implausible	Correct trials	3121	2622	2667	2350
	Implausible	Incorrect trials	3090	3489	2363	3239
	Plausible	All trials	2632	2542	2449	2523
	Plausible	Correct trials	2852	2354	2536	2423
	Plausible	Incorrect trials	2379	3172	2314	2890

EQ: equi-biased verb TR: transitive-biased verb A: ambiguous U: unambiguous

Verbal Stroop task

Participants were significantly less accurate in the incongruent (97% correct) than in the neutral condition (99% correct) ($t(47) = 4.66, p < .001$). Trials that were incorrect and trials in which naming latency was shorter than 200 ms or longer than 4000 ms were removed for RT analysis (5 trials were excluded). Remaining naming latencies that were 2.5 *SD* below or above a participant's mean latency for each condition were replaced with the 2.5 *SD* cutoff value (2.6% of the data were affected). Results showed that the participants were significantly slower to name target colors in the incongruent condition (761 ms) than in the neutral condition (635 ms) ($t(47) = 19.09, p < .001$).

Nonverbal Stroop task

The participants showed perfect accuracy in both the incongruent and neutral conditions. Trials in which response latency was shorter than 200 ms and longer than 4000 ms were removed for RT analysis (no trials were excluded). The remaining correct response latencies that were 2.5 *SD* below or above a participant's mean latency for each condition were replaced with the 2.5 *SD* cutoff value (3.3% of the data were affected).

Results showed that the participants were significantly slower to make responses in the incongruent (572 ms) than in the neutral condition (493 ms) ($t(47) = 13.02, p < .001$).

Relationship between executive control and syntactic recovery

The means, standard deviations, reliability estimates for the accuracy and RT measures are presented in Table 3.5a and Table 3.5b, respectively. The nonverbal executive control measure was not included in the accuracy analysis due to perfect accuracy by all participants in the nonverbal Stroop task. As in Experiment 1, reliability estimates were calculated using the random split-half correlation method and adjusted by the Spearman-Brown prophecy formula.

Table 3.5a. Experiment 3 – correlation across younger comprehenders: Means, standard deviations, and reliability estimates for accuracy measures.

Accuracy measure	Mean	SD	Reliability
Overall ambiguity effect	19%	21%	0.61
Ambiguity effect - implausible equi-biased	15%	26%	0.31
Ambiguity effect - implausible transitive -biased	19%	35%	0.49
Ambiguity effect - plausible equi-biased	24%	35%	0.11
Ambiguity effect - plausible transitive -biased	18%	28%	0.13
Verbal Stroop effect	2%	3%	0.52

Table 3.5b. Experiment 3 – correlation across younger comprehenders: Means, standard deviations, and reliability estimates for RT measures.

RT measure	Mean	SD	Reliability
Overall ambiguity effect	480	426	0.78
Ambiguity effect - implausible equi-biased	463	511	0.77
Ambiguity effect - implausible transitive-biased	498	397	0.63
Verbal Stroop effect	126	46	0.88
Nonverbal Stroop effect	79	42	0.64

For correlation analysis, observations with studentized t value greater than $|3|$ were identified as outliers and removed from analysis. In the correlations on the accuracy data, one observation was removed from the correlation between the overall syntactic recovery measure and the verbal executive control measure. In the correlations on the RT

data, the same participant was identified as an outlier in all of the correlations, and another participant was identified as an outlier in two correlations: (i) between the syntactic recovery measure for the transitive-biased verb condition and the verbal executive control measure, (ii) between the syntactic recovery measure for the transitive-biased verb condition and the nonverbal executive control measure. A summary of the correlation results is presented in Table 3.6.

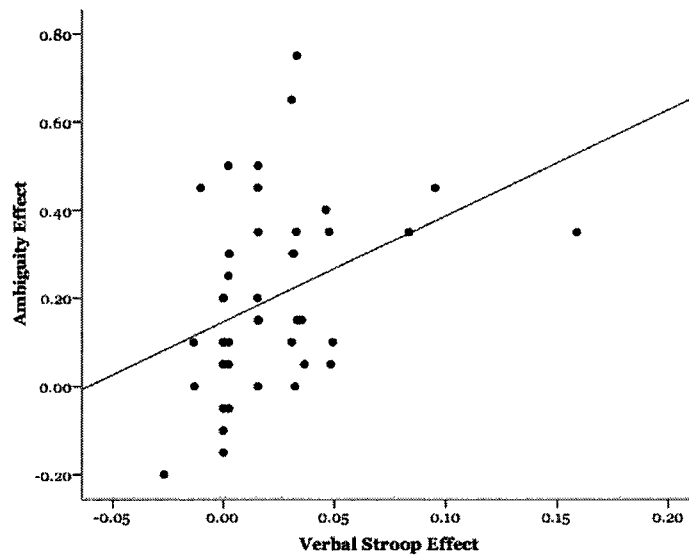
Table 3.6. Experiment 3 – correlation across younger comprehenders: Relationship between executive control and syntactic recovery.

Measure	Syntactic recovery	Executive control	
		Verbal	Nonverbal
Accuracy	Overall ambiguity effect	0.37*	
	Ambiguity effect - implausible equi-biased	0.20	
	Ambiguity effect - implausible transitive-biased	0.26†	
	Ambiguity effect - plausible equi-biased	0.15	
	Ambiguity effect - plausible transitive-biased	0.10	
RT	Overall ambiguity effect	0.03	-.12
	Ambiguity effect - equi-biased	0.05	-.16
	Ambiguity effect - transitive-biased	0.03	-.03

† indicates $p < .10$ * indicates $p < .05$

Consistent with the domain-specific executive control hypothesis, the verbal executive control accuracy measure positively and significantly correlated with the overall (collapsed across plausibility and verb bias) syntactic recovery accuracy measure across the participants ($r = .37, p = .01$) (see Figure 2). Within the implausible conditions, the correlation was positive and not significant in the less demanding, equi-biased verb condition ($r = .20, p = .17$), and positive and marginally significant in the more demanding, transitive-biased verb condition ($r = .26, p = .08$). The correlations within the plausible conditions were also positive but not significant ($p > .30$). No significant correlations were found between syntactic recovery and either of the executive control measure in the RT data (all $ps > .29$).

Figure 2. Experiment 3: Relationship between verbal Stroop effect and overall ambiguity effect on grammaticality judgment accuracy across participants.



Discussion of Experiment 3

Consistent with the domain-specific executive control hypothesis, participants with poorer verbal executive control were overall less successful in reinterpreting the structure of the garden path sentences. The prediction that syntactic garden path recovery would correlate with the verbal executive control measure *only* could not be tested in the accuracy measures due to perfect accuracy for the nonverbal executive control measure. (In the RT data, no correlations with either the verbal or nonverbal Stroop measures were significant.). Further study is needed to test the domain-general executive control hypothesis using a task in which some errors are made. As in Experiment 1, the correlations in the less demanding implausible conditions were positive but not significant. In contrast to Experiment 1, the correlations for the more demanding, plausible conditions were also not significant. As discussed previously, this may be due to the noise added into the syntactic recovery measure (i.e., due to a conflation of syntactic recovery and unrelated factors on the responses) in those conditions.

Summary of Experiment 1-3

The results in this set of experiments lend support for the domain-specific executive control hypothesis of garden path recovery. The verbal Stroop effect – a measure of verbal executive control, but not the nonverbal Stroop effect – a measure of nonverbal executive control, significantly correlated with the ambiguity effects – measures of garden path recovery, both at the semantic (Experiment 1) and syntactic level (Experiment 2). The findings suggest that domain-specific executive control is important for both semantic and syntactic reinterpretation of garden path sentences.

Chapter 3.

Deficit in Executive Control and Garden Path Reinterpretation

Previous patient case studies have shown an association between LIFG damage and executive control impairment (e.g., Barde, Schwartz, & Thompson-Schill, 2006; Hamilton & Martin, 2005, 2007; Thompson-Schill et al., 2002; see Jonides & Nee, 2006 for a review). One of the case studies was on a patient ML (Hamilton & Martin, 2005, 2007) who was among the patients tested in this work. Experiment 4 extended the examination of executive control ability in two new patients, another LIFG patient DW and a control patient LC. Experiment 5 examined the relationship between executive control deficit and semantic reinterpretation of garden path sentences while Experiment 6 examined the relationship between executive control deficit and syntactic reinterpretation of the sentences.

Patient Background

Patients ML and DW had lesions that included the LIFG area while patient LC had a non-LIFG (i.e., parietal) lesion. Each patient's single word processing, active and passive sentence comprehension, and short-term memory (STM) abilities were assessed as part of the patient screening process. A summary of the results is presented in Table 4.1. The single word and sentence processing abilities are reported in order to rule out the possibility that any patient difficulties in sentence processing might be attributed to difficulties with understanding single words or processing simple active and passive sentence structures. The STM measures are reported in order to provide further information on the relation between STM deficits and executive control deficits. In previous studies of patient ML, Hamilton and Martin (2005, 2007) hypothesized that a

deficit in the retention of semantic STM was due to difficulty in resolving interference.

The data from DW and LC will provide further information relevant to this hypothesis.

To assess single word comprehension, the Peabody Picture Vocabulary Test was used (choose one of four pictures that matches a spoken word) (PPVT: Dunn & Dunn, 1981). To assess single word production, the Philadelphia Picture Naming Test was used (name each single picture with a single word) (PPNT: Roach, Schwartz, Martin, Grewal, & Brecher, 1996). To assess comprehension of simple active and passive reversible sentences, a sentence-picture matching task was used (match each spoken sentence to an appropriate picture with distractor pictures showing a reversal of agent-patient roles or lexical substitutions). To assess short-term memory (STM), two tasks that tapped phonological retention – the forward digit span task (hear and repeat, in the original order, lists with varying number of digits) and the rhyme probe task (decide whether a probe word rhymes with any item in lists with varying number of items), and two tasks that tapped semantic retention – the synonymy judgment task (decide which two of three words are synonymous) and the category probe task (decide whether a probe word belongs to the same category as any item in lists with varying number of items) were used. To further assess STM capacities, each patient's score on each of the STM tasks was converted into a z-score (the scores of 8 other aphasic patients tested at Rice University and Temple University were included in this analysis). The patient's z-scores on the phonological STM tasks were added together to constitute the patient's composite phonological STM score, and the patient's z-scores on the semantic STM tasks were added together to constitute the patient's composite semantic STM score.

Table 4.1. Background of the patients tested in Experiment 4-6.

	Normal controls	Patient control	LIFG patients	
		LC	DW	ML
Age	varied	62	53	69
Sex	F, M	M	F	M
PPVT	100	120	73	107
PPNT	96%	96%	85%	98%
Sentence-picture matching	100%	92%	88%	95%
Phonological STM composite		0.11	0.28	-1.74
Forward digit span task	6.0 [†]	2.5	7.0	2.5
Rhyme probe task	7.5 ^{††}	5.0	3.3	1.8
Semantic STM composite		0.32	-0.61	-0.61
Synonym judgment task	†††	88%	83%	83%
Category probe task	6.1 ^{††}	2.5	1.8	1.5

PPVT: Peabody Picture Vocabulary Test

PPNT: Philadelphia Picture Naming Task

[†] Highest list length tested on normal controls, results reported in Martin, Lesch, and Bartha (1999).^{††} Results reported in Freedman and Martin (2001).^{†††} Results from normal controls were not available.

Patient ML. ML was a 68-year-old right-handed male, who suffered from a cerebrovascular accident in 1990 that resulted in a lesion comprising the LIFG, frontal areas more superior to the LIFG, and substantial areas of the left parietal lobe (see Figure 3). His single word processing was good (see Martin & Lesch, 1996); he performed above the mean for normal controls on both the PPVT (107 vs. normal mean = 100) (Dunn & Dunn, 1981) and the PPNT (98% correct vs. normal mean = 96% correct) (Roach et al., 1996). He also had good comprehension of simple active and passive sentences, performing at 95% correct on the sentence-picture matching task. ML received negative composite scores for both the phonological and semantic STM components (-1.74 for phonological and -.61 for semantic), suggesting that he was impaired in both the STM components relative to the other patients in the sample. Compared to previous reports (see Freedman & Martin, 2001; Martin & He, 2004; Martin & Lesch, 1996), his

semantic STM impairment was as severe as before (e.g., recent category probe span = 1.5 items vs. 1.8 items as reported in Freedman & Martin, 2001) while his phonological STM impairment had become more severe than before (e.g., recent rhyme probe span = 1.8 items vs. 3 items as reported in Freedman & Martin, 2001).

Patient DW. DW was a 53-year-old right-handed female, who suffered from a cerebrovascular accident in 2000 that resulted in a lesion comprising primarily the LIFG at the pars triangularis (BA 45), a small lesion in the left middle frontal gyrus, and an ischemic change in the right parietal lobe (see Figure 4). Her single word processing was fair, obtaining a score of 73 on the PPVT (Dunn & Dunn, 1981) and performing at 85% correct on the PPNT (Roach et al., 1996). Her comprehension of simple active and passive sentences was good, scoring 88% correct on the sentence-picture matching task. She received a positive composite score on the phonological STM component (.28) and a negative composite score on the semantic STM component (-.61), suggesting better phonological retention and worse semantic retention than the other patients in the sample.

Patient LC. LC was a 62-year-old right-handed male who suffered from a cerebrovascular accident in 2004 that resulted in a lesion in the left parietal lobe (image unavailable, details obtained from the patient's medical records). His single word processing was good, scoring 120 on the PPVT (Dunn & Dunn, 1981) and performing at 96% correct on the PPNT (Roach et al., 1996). He also had good comprehension of simple active and passive sentences, performing at 92% correct on the sentence-picture matching task. He received positive composite scores for both the phonological (.11) and semantic (.32) components, suggesting better retention on both the STM components compared to the other patients.

Overall, the background results show that the three patients could produce single words and understand single words and simple active and passive sentences adequately. Among them, ML had the worst phonological STM retention while DW was better than LC on the phonological STM component. For semantic STM retention, ML and DW were similarly impaired and worse than LC.

Figure 3. Structural brain scan of LIFG patient ML.

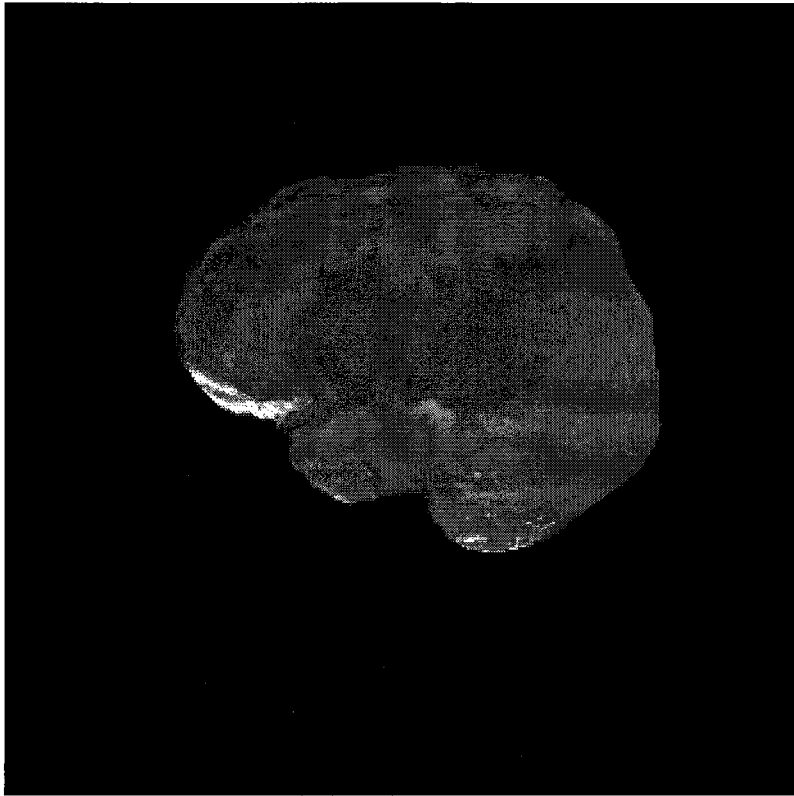
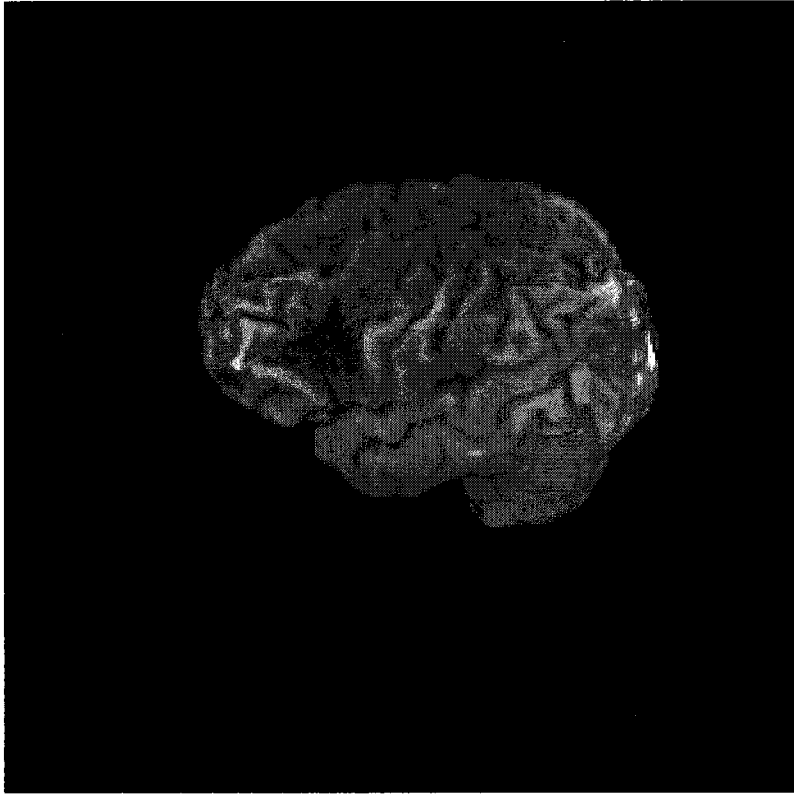


Figure 4. Structural brain scan of LIFG patient DW.



Experiment 4. LIFG Damage and Executive Control

Three tasks – the Stroop task (Stroop, 1935), the picture-word interference task (e.g., Schriefers, Meyer, & Levelt, 1990), and a modified recent-negatives task (Hamilton & Martin, 2007) – were used to examine executive control ability in the patients. These tasks induce interference (color word interfering with color ink naming, semantically related word distractors interfering with target picture naming, and semantically and phonological related list items interfering with recognition of probe items, respectively), thus requiring participants to focus attention on task-appropriate representations against distraction from task-inappropriate representations for correct performance. Patient's executive control ability was assessed by comparing the interference effect of each patient in each task against the interference effect of a normal control group. Exaggerated interference effects, defined as interference effects that were significantly greater than the normal mean interference effects under the Crawford and Howell's (1998) modified t-test procedure, reflect impaired executive control. Based on previous findings (e.g., Barde et al., 2006; Hamilton & Martin, 2005, 2007; Thompson-Schill et al., 2002), the LIFG patients DW and ML, but not the control patient LC, were expected to show consistently exaggerated interference effects across the three tasks.

Methods

Normal control participants. The participants were selected from a pool of older adults recruited from the Houston community. Their age ranged from 53 to 79 years old and they had attended at least some college. Not all of the normal controls were available for testing in all the tasks.

Stroop task. . The materials and the procedure of this task were the same as those in Experiment 1. Interference was defined as the mean difference in naming accuracy and latency between incongruent and neutral trials (neutral minus incongruent for accuracy and incongruent minus neutral for latency). The data for patient ML and 10 normal controls have been reported in Hamilton and Martin (2005).

Picture-word interference task. Participants saw pictures on the computer screen that included a word distractor simultaneously presented in each picture. They named the picture and ignored the word. Naming latency was recorded with a voice key. In the related condition (80 trials), the word distractor was semantically related to the picture name. In the unrelated condition (80 trials), the word distractor was unrelated to the picture name. The order of the trials was randomized for each participant. Interference was defined as the mean difference in naming accuracy and latency between related and unrelated trials (unrelated minus related for accuracy and related minus unrelated for latency). Fourteen normal controls performed this task.

Recent-negatives task. Participants saw lists of three words presented serially on the computer screen and a probe word presented after each list. The list words were presented for 1000 ms each, with a between-word interval of 100 ms. The probe word was presented for 750 ms, with an interval of 1100 ms between the list and the probe word. Participants indicated as quickly as possible whether the probe appeared in the list (*yes*) or not (*no*) by pressing the appropriate key. In the phonologically related and semantically related *same* list conditions (42 trials each), the probe was phonologically related and semantically related, respectively, to an item presented in the current list. In the phonologically related and semantically related *previous* list condition (42 trials

each), the probe was phonologically related and semantically related, respectively, to an item presented immediately before the current list. In the unrelated negative condition (168 trials), the probe was unrelated to currently and previously presented items. Besides the 336 negative trials, there were 336 positive trials in which the probe word was presented in the current list. The trials were presented in a fixed prerandomized order and were administered over two testing sessions. Interference was defined as the mean difference in response accuracy and latency in a related negative condition compared to the unrelated negative condition (unrelated negative minus related negative for accuracy, and related negative minus unrelated negative for latency). The data for patient ML and 14 normal controls have been reported in Hamilton and Martin (2007).

Results

The interference effects in the three tasks were first calculated for the normal controls as a group (significance was determined at the $p < .05$ level, two-tailed). The interference effects of each patient were then compared against the normal group's mean effects. An interference effect was considered as exaggerated if it was significantly greater than the normal mean using the Crawford and Howell's (1998) modified t-test procedure ($p < .05$, one-tailed). To account for generally longer response times in the patients, the comparison between the patient's and normal group's interference effects was done on both untransformed and transformed data (see Verhaeghen & De Meersman, 1998). For the untransformed RT analyses, data points that were 2.5 *SDs* beyond the mean of each condition in each subject were replaced with the 2.5 *SD* cutoff value. For the transformed RT analyses, all data (untrimmed) were included and log-transformed prior to analysis. A summary of the results is provided in Table 4.2.

Table 4.2. Experiment 4 – executive control: Mean interference effects (and standard deviations) for normal controls and patients.

	Normal controls	Patient control LC	LIFG patients DW ML	
Stroop interference effect				
Accuracy	6.3 (6.4)	3.6	36.0 [▲]	2.0
RT	197 (62)	553	966**	979**
Picture-word interference effect				
Accuracy	2.7 (2.8)	0.0	3.8	0.0
RT	35 (54)	-48	570**	623**
Recent-negatives interference effect				
Accuracy				
Phonologically related same list	1.9 (3.0)	-1.2	0.6	10.7 [▲]
Phonologically related previous list	0.3 (1.4)	-3.6	0.6	3.6 [▲]
Semantically related same list	0.0 (1.5)	6.0 [▲]	0.6	-1.2
Semantically related previous list	-0.2 (1.6)	-1.2	-1.8	1.2
RT				
Phonologically related same list	77 (51)	-43	168	387**
Phonologically related previous list	25 (70)	-34	186	276**
Semantically related same list	41 (38)	21	183**	356**
Semantically related previous list	5 (60)	-21	103	494**

[▲] Interference effects in accuracy that were significantly greater than normal group's means.

** Interference effects in RT that were significantly greater than normal group's means in both untransformed and log-transformed data.

Stroop task. As reported in Hamilton and Martin (2005), the normal group showed a significant interference effect of 6.3% ($SD = 6.4$) on naming accuracy ($t(9) = 3.09, p = .01$). The control patient LC showed a 3.6% effect, the LIFG patient DW a 36% effect, and the LIFG patient ML a 2% effect. While LC's effect and ML's effect were within 1 standard deviation away from the normal mean ($z = -.42$ and $-.67$ for LC and ML, respectively) and not significantly different from the normal mean (both $ts < 1$), DW's effect was 4.6 standard deviations above and significantly greater than the normal mean ($t(9) = 4.40, p < .001$).

For naming latency, the normal group showed a significant interference effect of 197 ms ($SD = 62$) ($t(9) = -10.0, p < .0001$). LC showed a 553 ms effect, DW a 966 ms effect, and ML a 979 ms effect. While LC's effect was 5.74 standard deviations above

and significantly greater than the normal mean in the untransformed data ($t(9) = 5.47, p < .001$), his effect was not significantly greater than the normal mean after log-transformation to correct for general slowing ($t(9) = 1.41, p = .10$). In contrast, both DW's effect and ML's effect were more than 12 standard deviations above the normal mean ($z = 12.40$ and 12.61 , respectively) and significantly greater than the normal mean in both untransformed ($t(9) = 11.82$ and 12.03 , respectively, both $ps < .001$) and log-transformed data ($t(9) = 1.89$ and 1.86 , respectively, both $ps < .05$).

Picture-word interference task. The normal group showed a significant interference effect of 2.7% ($SD = 2.8$) on naming accuracy ($t(13) = 3.61, p = .003$). LC showed a 0% effect, DW a 3.8% effect, and ML a reverse -1.2% effect, all of which were within 1.5 standard deviations from the normal mean and none were significantly greater than the normal mean (all three $ts < 1$).

For naming latency, the normal group showed a significant interference effect of 35 ms ($SD = 54$) ($t(13) = 2.43, p = .03$). LC showed a reverse -49 ms effect, DW a 570 ms effect, and ML a 623 ms effect. LC's effect was 1.55 standard deviations below the normal mean and not significantly different from the normal mean ($t(13) = -1.50, p = .16$, two-tailed). In contrast, both DW's effect and ML's effect were more than 10 standard deviations above the normal mean ($z = 10.00$ and 10.99 , respectively) and significantly greater than the normal mean in both untransformed ($t(13) = 9.66$ and 10.61 , respectively, both $ps < .001$) and log-transformed data ($t(13) = 3.16, p = .004$ and $t(13) = 3.47, p = .002$, respectively).

Modified recent-negatives task. The normal group showed a significant interference effect of 1.9% ($SD = 3.0$) on response accuracy for phonologically related

same list trials ($t(9) = 4.58, p = .001$). They did not show significant interference effects on response accuracy for the other three related conditions: .3% effect ($SD = 1.4$) for phonologically related previous list, 0% effect ($SD = 1.5$) for semantically related same list, and reverse -.2% effect ($SD = 1.6$) for semantically related previous list trials (all three $ts < 1$). LC showed small reverse effects in all the related conditions (see Table 4.2) (that were not significantly different from the normal means: $ts < 1$), except for the semantically related same list condition in which he showed a 6.0% interference effect, which was 3.84 standard deviations above and significantly greater than the normal mean ($t(13) = 3.71, p = .001$). DW showed small interference effects on accuracy (see Table 4.2) that were not significantly greater than the normal means in any of the conditions (all $ts < 1$). ML showed interference effects of 10.7% in the phonologically related same list and 3.6% in the phonologically related previous list condition, both of which were more than 2 standard deviations above ($z = 2.93$ and 2.25 , respectively) and significantly greater than the normal means ($t(13) = 2.83, p = .007$ and $t(13) = 2.18, p = .02$, respectively). For the two semantically related conditions, ML's accuracy effects (see Table 4.2) were not significantly greater than the normal means ($ts < 1$).

For response latency, the normal group showed a significant interference effect of 77 ms ($SD = 51$) in the phonological related same list condition ($t(13) = 5.57, p < .001$) and a significant interference effect of 41 ms ($SD = 38$) in the semantically related same list condition ($t(13) = 4.00, p = .002$). They showed a non-significant interference of 25 ms ($SD = 70$) in the phonological related previous list and a non-significant difference of 5 ms ($SD = 60$) in the semantically related previous list trials ($ts < 1.4$). LC did not show any interference effects on RT that were significantly greater than the normal means in

untransformed or log-transformed data (all t s < 1). DW showed interference effects on RT that were at least 1.64 SD s above the normal means across all four conditions in the untransformed data ($z = 1.78$ for phonologically related same list, $z = 2.30$ for phonologically related previous list, $z = 3.74$ for semantically related same list, and $z = 1.64$ for semantically related previous list). Among them, her phonologically related previous list and semantically related same list effects were significantly greater than the normal mean effects in the untransformed data ($t(13) = 2.23, p = .02$ and $t(13) = 3.62, p = .002$, respectively). After log transformation, only the semantically related same list effect remained significantly greater than the normal mean ($t(13) = 2.20, p = .02$). ML showed interference effects on RT that were at least more than 3.5 standard deviations above the normal means across all four conditions in the untransformed data ($z = 6.04$ for phonologically related same list, $z = 3.59$ for phonologically related previous list, $z = 8.29$ for semantically related same list, and $z = 8.18$ for semantically related previous list). All of the four effects were significantly greater than the normal means in both untransformed and log-transformed data (all p s < .03).

Discussion of Experiment 4

In line with previous findings (Barde et al., 2006, Hamilton & Martin, 2005, 2007; Thompson-Schill et al., 2002), the LIFG patients ML and DW consistently showed exaggerated interference effects across all three executive control tasks while the non-LIFG patient LC did not. These results added further evidence to the hypothesis that the LIFG serves to resolve interference in working memory (Jonides & Nee, 2006).

Hamilton and Martin (2005, 2007) have proposed that impaired performance in semantic STM tasks is caused by impairment in executive control, particularly by

impairment in the ability to inhibit irrelevant representations in working memory. If this account is correct, patients who are impaired in semantic STM tasks should be impaired in resolving interference in executive control tasks. To the extent that ML and DW both had impaired semantic STM ability and both showed exaggerated interference effects in the executive control tasks, the results are consistent with the Hamilton and Martin's view (but see Barde et al., 2006). However, it should be noted that LC's semantic STM composite was only slightly higher than that of ML or DW but he showed substantially reduced interference effects relative to these patients. Data from further patients would be needed to determine the extent to which semantic STM deficits and deficits in the resolution of interference in working memory are dissociable.

Experiment 5. LIFG Damage and Semantic Recovery

Experiment 1 showed that the ability to reinterpret the meaning of garden path sentences is related to (verbal) executive control ability. Experiment 4 showed that patients with damage to the LIFG area have impaired executive control. This experiment was carried out to examine whether LIFG damage, hence impaired executive control, is associated with impaired semantic reinterpretation of garden path sentences. The testing procedure of this experiment was similar to that of Experiment 2. Patients and normal controls read object/subject garden path and unambiguous control sentences, and answered for each sentence a yes/no question about the object interpretation of the ambiguous noun phrase. To increase the number of items per condition, only the sentences with later mismatching semantic information (the implausible sentences) were included in this experiment. The materials therefore consisted of two manipulated factors

– Verb Bias and Ambiguity. Each participant was tested on all of the conditions and items.

Predictions for the normal group

Experiment 2 has shown that, when the task focuses on meaning interpretation, the unambiguous structural marker (the comma) had no effect on final sentence meaning interpretation (the younger comprehenders were on average 69.3% correct on the ambiguous sentences and 69.8% correct on the unambiguous sentences). However, the reading time results indicate that the participants were sensitive to the ambiguity manipulation and experienced increased processing difficulty at the syntactic error cue (main verb region) in the ambiguous sentences. The processing difficulty was further found to be more pronounced for transitive-biased than for equi-biased verbs, suggesting a greater commitment to the initial object interpretation for transitive-biased verbs although the participants were equally likely to eventually arrive at the correct interpretation across the two verb bias conditions.

Prior aging research has suggested that older adults may have deficits in executive control (e.g., Hasher & Zacks, 1988). However, verbal Stroop results from the pre-testing of the normal (older) controls in this experiment showed a similar performance of this group to the younger participants in Experiment 1 (see the *Participants* section for more details). Given those results, the normal controls were expected to perform similarly to the younger participants on this garden path comprehension task. Specifically, there should be no difference among the ambiguity conditions in terms of sentence comprehension accuracy. However, there should be a significant main effect of

ambiguity and a significant interaction between ambiguity and verb bias on reading time at the main verb or the final region of the sentence.

Predictions for the patients

According to the executive control hypothesis, sentences with transitive-biased verbs are more difficult than those with equi-biased verbs because the transitivity bias toward the initial interpretation induces a greater degree of interference on the reinterpretation process. Given the LIFG patient's executive control deficit (i.e., interference resolution deficit), it was expected that the LIFG patients DW and ML would show an increased likelihood of maintaining the initial interpretation (leading to incorrect comprehension responses) in the more interfering, transitive-biased verb conditions. If the garden path sentences used in this experiment are interpreted similarly with and without the comma, as suggested by the results of the younger comprehenders in Experiment 2, then the LIFG patients should be impaired on both the ambiguous and unambiguous sentences that contain transitive-biased verbs. In other words, compared to the normal control group, the LIFG patients may show an exaggerated main effect of verb bias on comprehension accuracy.

Experiment 1 shows that comprehenders with poorer executive control may experience greater ambiguity effects on RT (as shown by the significant positive correlation between (verbal) executive control and ambiguity effect on RT in the plausible equi-biased verb condition). However, those with poorer executive control may also conversely show *reduced* ambiguity RT effects (as shown by the significant negative correlation between (verbal) executive control and ambiguity effect on RT in the plausible transitive-biased verb condition). Given those results, LIFG patients may be

expected to show either exaggerated increases or exaggerated decreases in ambiguity RT effects, compared to the normal group.

Given that the LIFG patients and the control patient LC did not score at a normal level in other language-related functions (e.g., reduced phonological and semantic short-term memory compared to normal controls), a general decrement in sentence processing might be expected as well. However, given that LC had normal executive control, the executive control hypothesis would predict that he would not show the specific exaggerated (or reduced) effects that were expected for the LIFG patients.

The LIFG area, also known as Broca's area, has traditionally been considered as an important area for syntactic functioning (general and specific syntactic processing factors have been hypothesized) (e.g., Berndt & Caramazza, 1980; Grodzinsky, 1990, 1995; Grodzinsky & Friederici, 2006; though see Dick et al., 2001, Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004 for negative evidence and see Martin, 2006 for a review). The structure of the garden path sentences used in this experiment follows a straightforward linear SVO word order that presumably does not involve complex syntactic operations (e.g., *While the man coached, the woman attended the party by herself*). Still, we might consider a simple account in which damage to the LIFG affects general syntactic processing, which in some way causes difficulty with syntactic revision that in turn leads to a deficit in semantic reinterpretation. This simple account would predict that the LIFG patients would perform at chance (50%) on all garden path sentences and correctly on the corresponding unambiguous versions (assuming that they can appreciate the significance of the comma for syntactic parsing).

Methods

Participants

The LIFG patients ML and DW and the control patient LC who participated in Experiment 4 were tested in this experiment. A new group of normal older controls (5 females and 3 males, mean age = 63.8, $SD = 6.4$) also participated in the experiment. Each participant was tested in two sessions, which were approximately one week apart. They received monetary compensation at a rate of \$10 per hour.

The new normal controls were additionally tested on the verbal Stroop task. Their mean Stroop interference effect and standard deviation on naming accuracy and latency is presented in Table 5.1a (the results of the normal control group from Experiment 4 are also included for comparison). As the table shows, the results of the normal controls tested in this experiment were similar to the results of those tested in Experiment 4.

Table 5.1a. Experiment 5 – executive control: Mean verbal Stroop effects (and standard deviations) for normal controls in Experiment 5-6 vs. for normal controls in Experiment 4.

	Accuracy	RT	
		Trimmed untransformed	Log-transformed
Controls - Experiment 5-6 (n = 8)	5.0% (4.3%)	187 (111)	.09 (.05)
Controls - Experiment 4 (n=10)	6.3% (6.4%)	197 (62)	.09 (.04)

As noted previously, prior research has suggested that ability to focus attention against interference from irrelevant information in working memory may be less efficient in older than in younger adults (e.g., Hasher, Lustig, & Zacks, 2007; Hasher & Zacks, 1988; but see, e.g., Verhaeghen & Cerella, 2002; Verhaeghen & De Meersman, 1998). To examine whether the normal (older) controls tested in this experiment had less efficient interference resolution mechanisms compared to the younger comprehenders tested previously, the Stroop performance of the normal controls in this experiment (n = 8) was compared to that of the younger comprehenders run in Experiment 1 (mean age = 19.5,

$SD = 1.3$, $n = 48$) in a mixed 2 (Condition: neutral vs. incongruent) x 2 (Group: old vs. young) ANOVA with Condition being a within-subject factor and Group a between-subject factor. The results are summarized in Table 5.1b.

Table 5.1b. Experiment 5 – executive control: Mean verbal Stroop effects (and standard errors) on naming accuracy and RT by condition for younger comprehenders in Experiment 1 vs. for normal (older) controls in Experiment 5-6.

Group	Incongruent	Neutral	$F_{1,54}$	Accuracy	RT	Log-RT
Young	99.2% (.2%)	96.5% (.6%)	Condition	27.53**	183.97**	221.30**
Old	98.3% (.5%)	93.3% (1.4%)	Group	5.58*	33.08**	29.68**
			C x G	2.45	2.84†	<1
Young	780 (15)	634 (10)				
Old	980 (15)	792 (25)				

C x G: condition x group † indicates $p < .10$ * indicates $p < .05$ ** indicates $p < .01$

As is clear from the table, the normal (older) controls were generally slower and made more errors than the younger participants (mean difference in RT = 179 ms, mean difference in accuracy = 2%, collapsing across conditions). However, the difference in terms of Stroop effects across the age groups was not significant for accuracy (5% vs. 3% effect, respectively, $p = .12$). For naming latency, the difference was marginally significant in untransformed RTs (187 ms for older vs. 146 ms for younger, respectively, $p = .10$), but after the data were log-transformed to account for general slowing in the older group, the Condition x Group interaction was not close to significance: $F < 1$. Thus, there was no evidence that the normal controls in this experiment had reduced executive control ability. The difference between these findings and those from earlier studies may be due to the fact that the older subjects in this experiment were younger than many older subjects tested in prior aging studies (e.g., Spieler, Balota, & Faust, 1996).

Materials

The materials were drawn from those used in Experiment 2 and included only the implausible items (see examples below, see also Appendix A for all experimental

sentences and Appendix B for examples of filler sentences). There were two counterbalancing lists of materials in which Ambiguity and Verb Bias were crossed. The number of filler sentences was doubled in this experiment to accommodate the within-subject design. Each participant was run on both lists. The order of list presentation was counterbalanced for the normal controls.

Equi-biased Verb (Implausible)

Ambiguous: While the man/ coached / the woman/ attended/ the party by herself.

Unambiguous: While the man/ coached, / the woman/ attended/ the party by herself.

Did the man coach the woman?

Transitive-biased Verb (Implausible)

Ambiguous. While the uncle/ visited / the child/ was/ missing him at home.

Unambiguous: While the uncle/ visited, / the child/ was/ missing him at home.

Did the uncle visit the child?

Filler sentences

While the father/ urged/ the daughter/ the son/ prepared/ breakfast in the kitchen.

Did the daughter prepare breakfast?

While the kite/ flew,/ the girl/ imagined/ herself riding away on it.

Did the girl want to ride the kite?

While the candidates/ debated/ the plan/ the moderator/ tried/ to stop them in vain.

Did the candidates agree on the plan?

Procedure

The procedure was similar to that of Experiment 2. Participants read the sentences one region at a time and pressed a button to proceed to the next region. The duration of time between the two button presses was recorded as the reading time for the region. Following the end of each sentence, a yes/no comprehension question appeared on the screen and participants answered by pressing a “yes” or “no” button. The order of the trials was pre-randomized and fixed for all participants.

Results

A summary of the overall accuracy for the experimental and filler sentences for the patients and normal controls is presented in Table 5.2. As the table shows, all three patients performed below the normal means on both the experimental and filler items. However, the patient's relatively high accuracy on the filler items (in the 80% range) indicates that they could perform the task adequately.

Table 5.2. Experiment 5 – overall sentence comprehension: Mean accuracy (and standard deviations) on experimental and filler sentences for normal controls and patients.

Subject	Sentences	
	Experimental	Filler
Normal controls	90% (7.8)	95% (1.7)
Control patient LC	61%	89%
LIFG patient DW	49%	82%
LIFG patient ML	51%	81%

The low accuracy rate of the patients on the experimental items led to problems with missing data in reading time analysis if only correct trials were to be included. Since the patients' processing of the sentences was of interest even when they arrived at the wrong answer, an examination of the patient's reading time data was done with both correct and incorrect trials included. To facilitate comparisons of patient vs. normal control performance, analysis of the normal control group's RT data also followed the same procedure (analyses with only correct reading times included yielded the same pattern of results).

Normal control group

The accuracy and reading time data of the normal controls were separately examined in a 2 (Ambiguity: ambiguous vs. unambiguous) x 2 (Verb Bias: neutral vs. transitive-biased) within-subject ANOVA in the subject analysis and a 2 x 2 mixed ANOVA in the item analysis (Verb Bias was a between-item factor and Ambiguity was a

within-item factor). As in the previous experiments, data from the syntactic error cue, main verb region and the semantic error cue, final region were included in the reading time analysis. Within each region, correct reading times that were 2.5 *SDs* beyond a participant's mean for each experimental condition were replaced with the 2.5 *SD* cutoff value (3.0% and 2.5% of the data were affected for the main verb and final region, respectively).

Comprehension accuracy. Overall, the normal controls averaged 90% correct on the experimental sentences, which was comparable to the performance of the younger participants in Experiment 1 (92% correct, implausible conditions only) and Experiment 2 (88% correct, implausible conditions only). Their means and standard errors for each condition are presented in Table 5.3a, and their *F*-statistics from the ANOVA results are presented in Table 5.3b.

Table 5.3a. Experiment 5 – garden path comprehension: Means (and standard errors) by condition for normal controls.

Condition	Accuracy	Main verb RT	Final RT
EQ ambiguous	89% (4%)	1199 (180)	2206 (209)
EQ unambiguous	88% (3%)	978 (131)	2135 (199)
TR ambiguous	92% (2%)	1182 (147)	2163 (268)
TR unambiguous	91% (3%)	943 (122)	1984 (262)

EQ: equi-biased verb

TR: transitive-biased verb

Table 5.3b. Experiment 5 – garden path comprehension: *F*-statistics for normal controls.

Variable	Accuracy		Main verb RT		Final RT	
	F _{1,7}	F _{2,38}	F _{1,7}	F _{2,38}	F _{1,7}	F _{2,38}
Verb Bias	4.44†	<1	<1	<1	1.32	<1
Ambiguity	<1	<1	10.83*	56.91**	2.24	6.81*
Ambiguity x Verb Bias	<1	<1	<1	<1	<1	1.03

† indicates $p < .10$

* indicates $p < .05$

** indicates $p < .01$

Similar to the younger comprehenders in Experiment 2 (in which the same yes/no testing procedure was used), the normal controls showed no effect of ambiguity on comprehension accuracy. Specifically, they were as likely to reject the initial object

interpretation in the ambiguous (90% correct) as in the unambiguous conditions (89% correct) (main effect of Ambiguity: $F_s < 1$). Collapsing across ambiguity conditions, they showed a small reverse verb bias effect, performing slightly though not significantly less accurately on the equi-biased (88% correct) vs. transitive-biased verb items (92% correct; main effect of Verb Bias: $F_1(1, 7) = 4.44, p = .07, F_2 < 1$). They showed no difference in the effect of ambiguity across the verb bias conditions (Verb Bias x Ambiguity interaction: $F_s < 1$).

Reading time. The reading time results at the main verb region confirmed that the ambiguity manipulation was successful in leading the normal controls into temporarily adopting the object interpretation. They were significantly slower to read the syntactic error cue at the main verb of the ambiguous (1190 ms) vs. unambiguous sentences (960 ms) (main effect of Ambiguity: ($F_1(1, 7) = 10.83, p = .01, F_2(1, 38) = 56.91, p < .001$). However, unlike the younger comprehenders in Experiment 2 who showed a significantly larger ambiguity effect in the transitive-biased vs. equi-biased verb condition (522 ms and 311 ms, respectively), the ambiguity effect of the normal controls in this experiment was slightly and not significantly larger for the transitive-biased (239 ms) vs. equi-biased verb condition (221 ms) (Ambiguity x Verb Bias interaction: $F_s < 1$). This suggests that the normal controls were similarly committed to the initial interpretation across the verb bias conditions and consequently experienced a similar degree of processing difficulty across the two conditions upon encountering the syntactic error cue. Also unlike the young comprehenders in Experiment 2, whose reading was slower in the transitive-biased verb than in the equi-biased verb condition, the normal controls in this experiment

showed no such reliable slowing in their reading (1190 ms vs. 1062 ms, respectively; main effect of Verb Bias: $F_s < 1$).

At the final region where the semantic error cue was located, the normal controls continued to experience some degree of increased processing difficulty in the ambiguous (2185 ms) vs. unambiguous condition (2059 ms). The main effect of Ambiguity was significant by items ($F_2(1, 38) = 6.81, p = .01$) but not significant by subjects ($F_1(1, 7) = 2.24, p = .18$). The ambiguity effect was larger for the items containing transitive-biased verbs (180 ms) than for those containing equi-biased verbs (71 ms). However, the interaction of Verb Bias and Ambiguity was not significant ($F_s < 1.1$). Again, collapsing across the ambiguity conditions, the normal controls read the final region of the sentences across the two verb bias types at a similar speed (main effect of Verb Bias: $F_1 = 1.31, F_2 < 1$). (Comparison of the results of the normal controls to those of the younger comprehenders is difficult for this region because there was an additional manipulation of Plausibility (plausible vs. implausible) in the experiments with the younger comprehenders.)

Patients

A mixed 2 x 2 ANOVA was performed separately for each patient on the accuracy vs. reading time data with items serving as a random factor (Verb Bias was a between-item factor and Ambiguity was a within-item factor). A summary of the accuracy results is presented in Table 5.4 and a summary of the reading time results is presented in Table 5.5a and 5.5b.

Comprehension accuracy

Table 5.4. Experiment 5 – garden path comprehension: Mean accuracy (and standard errors) by condition for normal controls and patients, and F -statistics from item analysis for patients.

Cond	Normals	LC	DW	ML	$F_{1,38}$	LC	DW	ML
EQ A	89% (4)	50% (11)	60% (11)	50% (11)	Verb Bias	<1	7.06*	<1
EQ U	88% (3)	60% (10)	70% (11)	55% (11)	Ambiguity	4.03†	1.92	<1
TR A	92% (2)	55% (11)	25% (11)	45% (11)	B x A	<1	<1	<1
TR U	91% (3)	80% (10)	40% (11)	55% (11)				

EQ: equi-biased TR: transitive-biased † indicates $p < .1$ * indicates $p < .05$
A: ambiguous U: unambiguous

Patient LC. Overall LC was much less accurate than the normal controls, averaging at 61% correct on the experimental sentences (normal mean = 90% correct, $SD = 7.8$). He showed a marginally significant effect of ambiguity on accuracy, performing less accurately on the ambiguous (53% correct) than on the unambiguous sentences (70% correct) (main effect of Ambiguity: $F(1, 38) = 4.03, p = .052$). His ambiguity effect was larger for the sentences containing transitive-biased verbs (25%) than for those containing equi-biased verbs (10%), though the interaction between Verb bias and Ambiguity was not significant ($F < 1$). Finally, similar to the normal controls, LC showed a similar reverse pattern of verb bias effect, performing more accurately on the sentences containing transitive-biased verbs (68% correct) than on those containing equi-biased verbs (55% correct). However, his main effect of Verb Bias was not significant ($F < 1$).

Patient DW. Overall DW was less accurate than both the normal controls and LC, averaging only 49% correct on the experimental sentences. However, in contrast to the normal controls and LC and as expected, she was significantly less accurate on the sentences containing transitive-biased verbs than on those containing equi-biased verbs (33% vs. 65% correct, respectively; main effect of Verb Bias: $F(1, 38) = 7.06, p = .01$). Similar to LC, she was generally less accurate on the ambiguous sentences (43% correct, compared to 55% correct on the unambiguous sentences). When broken down by verb

bias, her ambiguity effect was also larger for the sentences containing transitive-biased verbs (15%) than for those containing equi-biased verbs (10%). However, her main effect of Ambiguity was not reliable ($F(1, 38) = 1.92, p = .17$), nor was the interaction between Verb Bias and Ambiguity ($F < 1$).

To test whether DW's verb bias effect differed from that of LC's, a 2 (Patient Type: LIFG vs. non-LIFG) x 2 (Verb Bias: equi-biased vs. transitive-biased) between-subject ANOVA was performed on the patient's percentage accuracy. Results showed a significant interaction between Patient Type and Verb Bias ($F(1, 76) = 6.54, p = .01$) (neither the main effect of Patient Type nor of Verb Bias was significant, $F < 2.1$). Two further independent samples t-tests showed that DW was as accurate as LC on the sentences containing equi-biased verbs ($t < 1$) but significantly less accurate than LC on the sentences containing transitive-biased verbs ($t(38) = -2.97, p = .005$).

Patient ML. ML was also less accurate than the normal controls and LC, showing an overall accuracy rate of 51% correct. However, different from LC – who performed better on the unambiguous sentences, and DW – who performed better on the sentences containing (less interfering) equi-biased verbs, ML showed little variation across conditions and performed essentially around chance level (50%) on all conditions. Given that he had the worst phonological short-term memory deficit compared to the other patients and given the memory demand of the phrase-by-phrase self-paced reading task, one may question if he could do the task at all. The results on the filler sentences, which were as at least as long and complex as the experimental sentences (e.g., sentence: *While the father urged the daughter the son prepared breakfast in the kitchen*, question: *Did the daughter prepare breakfast?*), suggests that he could. As shown in Table 5.2, his overall

accuracy rate on the filler items was good and comparable to that of the other patients (ML – 81% correct, compared to LC – 88% correct and DW – 82% correct).

The ANOVA results on the experimental sentences showed that none of the effects were significant for ML (main effect of Ambiguity and of Verb Bias, and interaction between Verb Bias and Ambiguity: all F s < 1). However, his numerical pattern of ambiguity effects was similar to that of the other patients. Collapsed across verb bias conditions, he was less accurate on the ambiguous (48% correct) vs. unambiguous sentences (55% correct). His ambiguity effect was larger for the sentences containing transitive-biased verbs (10%) than for those containing equi-biased verbs (5%). Similar to DW, collapsing across ambiguity conditions, he was slightly less accurate on the transitive-biased (50% correct) than on the equi-biased verb items (53% correct).

Comparison against the normal group. Similar to Experiment 4, the effects of each patient were compared against the normal control mean effects, using the Crawford and Howell's (1998) modified t-test procedure. Performance was considered impaired if the patient showed an accuracy effect that was significantly above the normal mean (significance level at $p < .05$, one-tailed, unless specified otherwise). Compared to the normal group on the overall ambiguity effect (normal mean = $-.9\%$, $SD = 4.0$), LC's effect of 17.5% ($z = 4.62$), DW's effect of 12.5% ($z = 3.36$) and ML's effect of 7.5% ($z = 2.11$) were all significantly greater than the normal mean (all $ps < .05$). Within the equi-biased verb condition, none of the patient's effects were significantly greater than the normal mean (10% effect for LC and DW and 5% effect for ML, compared to normal mean = -1.3% , $SD = 7.44$; all $ps \geq .10$). Within the transitive-biased verb condition, the

ambiguity effects of all three patients were significantly greater than the normal mean (LC showed a 25% effect, $z = 5.17$, DW showed a 15% effect, $z = 3.15$, ML showed a 10% effect, $z = 2.14$, compared to normal mean = -. 6%, $SD = 5.0$; all $ps < .05$).

For the main effect of verb bias, similar to the normal group (normal mean = - 3.4%, $SD = 4.62$), LC showed a reverse verb bias effect of -12.5%, which was 1.96 standard deviations below but not significantly different from the normal mean ($t(7) = -1.85$, $p = .11$, two-tailed). Importantly, DW's verb bias effect of 32.5% was in the theoretically predicted direction and far greater than the normal mean (7.34 standard deviations above the normal mean, $t(7) = 7.34$, $p < .001$). ML's verb bias effect of 2.5% was also in the theoretically predicted direction, though not significantly greater than the normal mean (1.29 standard deviations above the normal mean, $t(7) = 1.21$, $p = .13$).

Further verification from reading time

As with the normal control data, data from the syntactic disambiguating main verb and the semantic disambiguating final region were included in the reading time analysis for the patients. To account for generally longer reading times in the patients, both untransformed and log-transformed data were analyzed (Verhaeghen & De Meersman, 1998). In the untransformed reading time analysis, data points that were 2.5 SD s beyond the mean of each condition in each participant were replaced with the 2.5 SD cutoff value (less than 3% of each patient's main verb data and less than 4% of each patient's final region data were affected). All data (untrimmed) were included in the log-transformed reading time analysis. The results for the main verb region are summarized in Table 5.5a and the results for the final region are summarized in Table 5.5b.

The RT results of Experiment 1 demonstrates that those with poorer executive controls may exhibit patterns of either increased or decreased ambiguity effects on RT across different interfering conditions. Because it was difficult to predict a priori which of the patterns would apply to which of our executive control patients (who had varying degrees of executive control impairment), when comparing against the normal control group, a patient's RT performance was considered impaired if s/he showed an ambiguity effect that was significantly above or below the normal mean (significance level at $p < .05$, two-tailed, according to the Crawford & Howell's 1998 modified t-test procedure).

Table 5.5a. Experiment 5 – garden path comprehension: Mean reading times (and standard errors) by condition at the main verb region for normal controls and patients, and F -statistics from item analysis for patients.

Cond	Normals	LC	DW	ML	F(1,38)	LC	DW	ML
EQ A	1199 (180)	846 (55)	1105 (75)	2439 (307)	Verb Bias	<1	<1	<1
EQ U	978 (131)	773 (30)	1088 (65)	1962 (308)	Ambiguity	6.32*	<1	<1
TR A	1182 (147)	928 (55)	1121 (75)	2013 (307)	VB x A	<1	<1	<1
TR U	943 (122)	780 (30)	1114 (65)	1836 (308)				

EQ: equi-biased verb TR: transitive-biased verb * indicates $p < .05$
A: ambiguous U: unambiguous VB x A: Verb bias x Ambiguity

Main verb region. The results suggest that, in this task, LC was sensitive to the syntactic error cue while the LIFG patients DW and ML were not. LC showed an overall ambiguity effect of 110 ms that was significant in the ANOVA item analysis ($F(1, 38) = 6.32, p = .02$). In contrast, DW showed an overall ambiguity effect of 12 ms and ML showed an overall ambiguity effect of 327 ms that were not significant in the ANOVA item analysis ($F_s < 1$). Although none of the patients showed ambiguity effect sizes that significantly differed across the verb bias conditions (Ambiguity by Verb Bias interaction: all three $F_s < 1$), it is interesting to note that while LC showed numerically greater ambiguity effect in the more interfering transitive-biased verb condition (ambiguity effect = 73 ms for equi-biased vs. 148 ms for transitive-biased verbs), ML

showed the opposite pattern of decreased sensitivity to the syntactic error cue in the more interfering condition (ambiguity effect = 477 ms for equi-biased vs. 177 ms for transitive-biased verbs) (DW showed a pattern similar to that of ML, though her effects were quite small in both conditions, ambiguity effect = 17 ms for equi-biased vs. 7 ms for transitive-biased verbs).

Compared to the normal control group, who showed an overall ambiguity effect of 230 ms ($SD = 198$), all three patients showed effects that were within 1.5 standard deviations from the normal mean (again, LC's effect = 110 ms, DW's effect = 12 ms, and ML's effect = 327 ms). For verb-specific ambiguity effects, the normal group showed a 221 ms effect ($SD = 217$) for equi-biased verbs and a 239 ms effect ($SD = 185$ ms) for transitive-biased verbs. All three patients showed verb-specific ambiguity effects that were also within 1.5 standard deviations from the normal means. None of the patient's overall or verb-specific ambiguity effects were significantly different from the normal means in either the untransformed or log-transformed data (all $ps > .27$).

Table 5.5b. Experiment 5 – garden path comprehension: Mean reading times (and standard errors) by condition at the final region for normal controls and patients, and F -statistics from item analysis for patients.

Cond	Normals	LC	DW	ML	F(1,38)	LC	DW	ML
EQ A	2206 (209)	1553 (111)	3432 (366)	5593 (786)	Verb Bias	6.52*	<1	<1
EQ U	2135 (199)	1813 (137)	2822 (234)	7148 (824)	Ambiguity	1.77	3.02†	1.11
TR A	2163 (268)	1325 (111)	3197 (366)	5553 (786)	Bias x Am	<1	<1	<1
TR U	1984 (262)	1383 (137)	2925 (234)	5727 (824)				

EQ: equi-biased TR: transitive-biased

† indicates $p < .10$

* indicates $p < .05$

A: ambiguous U unambiguous

Final region. LC showed a significant reverse main effect of Verb Bias in the ANOVA item analysis, reading the equi-biased verb items significantly more slowly than reading the transitive-biased verb items (verb bias effect = -329 ms; main effect of Verb Bias: $F(1, 38) = 6.52$, $p = .02$, see Discussion for a possible explanation) (main effect of

Ambiguity, and Verb Bias x Ambiguity interaction effect were not significant: $F_s < 1.8$). Compared to the normal group, who also showed a reverse verb bias effect (mean = -95 ms, $SD = 236$), LC's reverse verb bias effect was within 1 standard deviation from the normal mean and not significantly different from the normal mean in either the untransformed or log-transformed data (both $ps > .31$).

In the ANOVA item analysis, DW showed a marginally significant main effect of ambiguity only (effect size = 441 ms, main effect of Ambiguity: $F(1, 38) = 3.02, p = .09$) (Verb bias x Ambiguity interaction and main effect of Verb Bias were not significant: $F_s < 1$). Within the equi-biased verb condition, her ambiguity effect of 609 ms was highly significant in the item analysis ($t(19) = 2.98, p = .008$). When compared to the 71 ms ambiguity effect for equi-biased verbs ($SD = 304$) in the normal controls, her effect was 1.77 standard deviations above the normal mean in the untransformed data ($t(7) = 1.67, p = .14$; after log-transformation: $t < 1$). In the more interfering, transitive-biased verb condition, DW's ambiguity effect of 272 ms was not significant in the ANOVA item analysis ($t < 1$), which, when compared to the 180 ms ambiguity effect ($SD = 270$ ms) in the normals, was within half a standard deviation from the normal mean and was not significantly greater than the normal mean in the untransformed or log-transformed data ($ts < 1$).

In the ANOVA item analysis, ML did not show any significant effects on reading time (main effect of Ambiguity: $F(1, 38) = 1.11, p = .30$; Verb Bias x Ambiguity interaction effect and main effect of Verb Bias: $F_s < 1$). Interestingly, he showed a pattern of hypo-sensitivity to ambiguity in this final region. He was numerically to read in the ambiguous vs. unambiguous conditions overall (reverse overall ambiguity effect =

-864 ms), which, when compared to an overall ambiguity effect of 128 ms ($SD = 234$) in the normal group, was 4.24 standard deviations below and significantly smaller than the normal control mean in the untransformed data ($t(7) = -4.00, p = .005$; after log-transformation: $t(7) = -1.41, p = .20$). Within each verb bias category, ML's ambiguity effect of -1555 ms for equi-biased verbs was 5.35 standard deviations below and significantly smaller than the normal control mean effect (of 71 ms) in the untransformed data ($t(7) = -5.04, p < .002$; after log-transformation: $t(7) = -1.66, p = .14$). His ambiguity effect of -174 ms in the transitive-biased verb condition was 1.31 standard deviations below the normal mean effect and not significantly smaller than the normal mean in either the untransformed or log-transformed data (both $ps > .25$).

Discussion of Experiment 5

As expected, the LIFG patient DW was significantly less accurate in the more interfering transitive-biased verb condition, which was true for both ambiguous *and* unambiguous sentences. Her verb bias effect on performance accuracy was significantly greater than the normal control mean and significantly greater than the effect found for the control patient LC. She also showed some evidence of an exaggerated ambiguity effect on RT at the final region of the equi-biased verb sentences. While DW showed impaired accuracy performance only in the more interfering condition (her accuracy performance was as good as that of the control patient LC in the less interfering, equi-biased verb condition), the other LIFG patient ML, who had a more severe interference resolution deficit, showed impaired accuracy performance in both the verb bias conditions. Interestingly, different from DW, he showed some evidence of an

exaggeratedly *reduced* ambiguity effect on RT at the semantic error cue region. Together, the current results provide good support for the executive control hypothesis.

Given that ML performed at chance in all the conditions, one may hypothesize that he had a general syntactic deficit that not only affected his processing of the garden path sentences but also of the unambiguous control sentences. However, numerically his ambiguity effect at the main verb was larger than the mean for normal controls, suggesting sensitivity to the comma cue. His high variability in RTs is no doubt the source of the non-significance of this effect. His good performance on the filler sentences also provides evidence against the general syntactic deficit explanation. The results of Experiment 6 will also help to confirm that he does not have a general syntactic processing deficit.

Overall all the patients including the control patient LC performed worse than the normal control group and showed ambiguity effects on comprehension accuracy that were significantly greater than the normal level. This suggests that the object/subject garden path sentences are generally difficult for the patients to comprehend. Patient LC also showed an unexpected pattern of reverse verb bias effects in several analyses that also appeared in the normal control group. An inspection of the experimental sentences shows that the equi-biased verb items contained more ambiguous noun phrases that were inanimate objects (e.g., *swept - the porch*, *drove - the minivan*, *wrote - the note*) than did the transitive-biased verb items. LC's (mostly numerical) reverse verb bias effect pattern might have been due to a bias toward taking inanimate objects as the direct object of the verb.

To sum up, the current results strongly suggest that executive control is critical for sentence comprehension. Patients with an executive control deficit may develop an incorrect understanding of a verbal message and continue to hold on to that misunderstanding despite contradicting evidence, if the message happens to contain words or phrases that have strong associations to structures or meanings not intended by the speaker or writer.

Experiment 6. LIFG Damage and Syntactic Recovery

Experiment 5 shows that LIFG damage is associated with severe meaning reinterpretation problems. This experiment was carried out to examine whether LIFG damage, hence executive control deficit, is associated with a deficit in the reinterpretation of syntactic structures of garden path sentences. The materials were the same as those used in Experiment 5 (also only implausible sentences were included). Participants judged the grammaticality of the sentences in this experiment.

Previous research has shown that impaired comprehension is not necessarily associated with impaired grammatical processing (e.g., Linebarger, Schwartz, & Saffran, 1983; see Martin, 2006 for a review). Patients may have a deficit in understanding the meaning of certain sentences (e.g., understanding who did what to whom in *The boy is followed by the girl*) and at the same time demonstrate good sensitivity to various aspects of grammar (e.g., have intact knowledge of verb's structural requirements and able to determine that a sentence containing an obligatorily intransitive verb used in the transitive form such as **The policeman was talking a woman* is ungrammatical). All three patients tested in this study were assessed on sentence processing during patient screening (see Patient Background) and had shown good comprehension performance

even on sentences with a non-canonical word order like reversible passive sentences (*The boy is followed by the girl*). None of the patients tested in this study therefore had a deficit in grammatical processing per se.

The claim to be tested in this experiment is that LIFG-based executive control patients would exhibit a grammatical processing deficit when aspects of that processing depend upon an intact ability to resolve interference in working memory. As argued previously, garden path sentences typically contain cues that support an incorrect alternative structure. Given that those cues create interference on the revision process, ability to resolve interference should be critical for a successful revision outcome. Experiment 3 has indeed shown that healthy younger individuals with poorer executive control were less able to revise the structure of object/subject garden path sentences (e.g., showing bigger ambiguity effects on grammaticality judgment accuracy). The LIFG-based executive control patients were therefore expected to show a specific deficit in syntactic revision under condition of strong interference from irrelevant cues.

It has been assumed in several previous experiments that garden path reinterpretation is specific to the ambiguous condition. Worse performance in the ambiguous compared to the unambiguous baseline condition (i.e., ambiguity effect) was taken as a measure of garden path reinterpretation ability. However, it is clear from the results of the previous experiments that the comma cue is not uniformly taken as a perfect marker of intransitivity - that is, the unambiguous condition was ineffective in preventing garden path misinterpretations. The much less than perfect *absolute* level of grammaticality judgment accuracy found in the unambiguous conditions of Experiment 3 (with younger comprehenders) proves the case. In the extreme condition in which the

comma cue is totally ineffective, performance across both ambiguity conditions would be similarly poor and there would be no difference in performance between the ambiguous and unambiguous condition. DW's semantic reinterpretation results in Experiment 5 provide an example of such an extreme case.

A deficit in syntactic revision may therefore be manifested in exaggerated ambiguity effects if the comma cue is at least relatively effective in precluding syntactic misinterpretations. Alternatively, the deficit may be manifested in exaggerated differences in performance between a less and a more interfering condition (e.g., conditions involving equi-biased vs. transitive-biased verbs) collapsing across ambiguity conditions (i.e., exaggerated verb bias effect). As in the previous experiments, deficits were defined relative to the normal control group's level of performance. The executive control hypothesis predicts that the LIFG patients ML and DW would show a deficit in at least the more interfering condition while the control patient LC would not. The general syntactic deficit hypothesis predicts that the LIFG patients would show impaired performance in all (both the less and more interfering) conditions.

Methods

Participants

The same participants from Experiment 5 participated in this experiment, which was separated from Experiment 5 by at least two weeks. Each participant was tested in two sessions, which were approximately one week apart. They received monetary compensation at a rate of \$10 per hour.

Materials

The experimental materials consisted of two counterbalancing lists that were the same as those in Experiment 5 (only the implausible items were included to increase the number of items per condition, see Appendix A for all experimental sentences). The filler sentences were similar to that in Experiment 2 (both grammatical and ungrammatical fillers were included, see Appendix C for examples), except that the number of fillers was doubled to accommodate the within-subject design of this experiment. Each participant was tested on both lists. The order of list presentation was counterbalanced for the normal controls.

Participants judged the grammaticality of the sentences after reading each sentence. An acceptance of grammaticality (a “yes” response) was the correct response for the experimental sentences. The “yes” response to the ambiguous sentences indicates that the participant has successfully revised the structure of the sentence.

Procedure

The procedure was similar to that of Experiment 5, except that in this experiment the participants pressed one of two buttons to indicate that the sentence was grammatical (“yes”) or ungrammatical (“no”).

Results

A summary of the overall accuracy on the experimental and filler sentences for the patients and normal controls is presented in Table 6.1. As the table shows, all three patients performed better on this grammaticality judgment task than on the comprehension task in Experiment 5, exhibiting performance that was much more comparable to that of the normal control group on both the experimental and filler sentences. For the filler sentences, both LC (85% correct) and DW (69% correct)

performed above the normal mean (65% correct) while ML's performance was only slightly below the normal mean level (60% correct). A similar result applied to the experimental sentences, with DW (81% correct) and LC (76% correct) performing above and ML (66% correct) performing below the normal mean (74% correct). As in Experiment 5, reading time data were analyzed for the normal controls and patients with both *correct and incorrect* trials included.

Table 6.1. Experiment 6 –overall grammaticality judgment: Mean accuracy (and standard deviations) on experimental and filler sentences for normal controls and patients.

Subject	Sentences	
	Experimental	Filler
Normal controls (n = 8)	74% (28%)	65% (20%)
Patient control LC	76%	85%
LIFG patient DW	81%	69%
LIFG patient ML	66%	60%

Normal control group

Judgment accuracy and reading time were examined in a 2 (Verb Bias: equi-biased vs. transitive-biased) x 2 (Ambiguity: ambiguous vs. unambiguous) within-subject ANOVA in the subject analysis and a 2 x 2 mixed ANOVA in the item analysis (Ambiguity was a within-item factor and Verb Bias was a between-item factor). The reading time analysis was on the main verb and the final region of the sentence. Within each region, reading times that were 2.5 *SD* beyond a participant's mean for each experimental condition were replaced with the 2.5 *SD* cutoff value (2.8% of the data for the main verb and 3.4% of the data for the final region were affected, respectively).

Grammaticality judgment accuracy. Overall, the normal controls were accurate on an average of 74% of the experimental trials, which was better than the 59% averaged accuracy rate of the young comprehenders in Experiment 3 (implausible sentences only).

A summary of the means and standard errors for each condition for the normal controls and a summary of the ANOVA results for the normal controls are presented in Table 6.2a and Table 6.2b, respectively.

Table 6.2a. Experiment 6 – garden path grammaticality judgment: Means (and standard errors) by condition for normal controls.

Condition	Accuracy	Main verb RT	Final RT
EQ ambiguous	72% (11%)	1168 (156)	2560 (333)
EQ unambiguous	79% (8%)	921 (114)	2760 (591)
TR ambiguous	72% (10%)	1144 (143)	2113 (326)
TR unambiguous	73% (11%)	928 (121)	2368 (414)

EQ: equi-biased verb TR: transitive-biased verb

Table 6.2b Experiment 6 – garden path grammaticality judgment: *F*-statistics for normal controls.

Variable	Accuracy		Main verb RT		Final RT	
	F _{1,7}	F _{2,38}	F _{1,7}	F _{2,38}	F _{1,7}	F _{2,38}
Ambiguity	2.70	4.82*	17.61**	71.22**	<1	1.88
Verb Bias	1.38	2.46	<1	<1	4.29†	4.97*
Ambiguity x Verb Bias	1.38	<1	<1	<1	<1	<1

† indicates $p < .1$

* indicates $p < .05$

** indicates $p < .05$

Unlike the younger comprehenders in Experiment 3, who showed a large ambiguity effect on grammaticality judgment accuracy (implausible sentences only: 50% correct on ambiguous vs. 68% correct on unambiguous sentences), the normal controls were only slightly less accurate on the ambiguous (72% correct) vs. unambiguous sentences (76% correct). The main effect of Ambiguity was significant by items ($F_1(1, 38) = 4.82, p = .03$) but not significant by subjects ($F_2(1, 7) = 2.70, p = .14$). Also unlike the younger comprehenders, who showed a pattern of a larger ambiguity effect for the transitive- vs. equi-biased verb items in Experiment 3 (implausible sentences only: 19% vs. 15% effect, respectively), the normal controls showed a reverse pattern of a slightly but not significantly larger ambiguity effect for the equi-biased (8% effect) vs. transitive-biased verb items (1% effect) (Verb Bias x Ambiguity interaction: $F_s < 1.4$). Collapsing

across ambiguity conditions, the normal controls were slightly but not significantly less accurate in the transitive-biased verb (73% correct) vs. equi-biased verb condition (76% correct) (main effect of Verb Bias: $F_s < 2.5$).

Reading time. The reading time results at the main verb region confirmed that the normal controls were, again, garden-pathed. Similar to the younger comprehenders in Experiment 3, they were significantly slower to read the main verb in the ambiguous (1156 ms) vs. unambiguous condition (924 ms) (main effect of Ambiguity: $F_1(1,7) = 17.61, p = .004, F_2(1,38) = 71.22, p < .001$). Their ambiguity effect was slightly but not significantly larger for the equi-biased (247 ms) vs. transitive-biased verb items (217 ms) (Verb bias x Ambiguity interaction: $F_s < 1$). Their overall reading time was similar across the verb bias conditions (1044 ms in the equi-biased verb vs. 1036 ms in the transitive-biased verb condition) (main effect of Verb Bias: $F_s < 1$).

The reading time pattern at the final region of the normal controls was generally similar to that of the younger comprehenders in Experiment 3. They showed a marginal main effect of Verb Bias on reading time, slower on the equi-biased (2660 ms) than on the transitive-biased verb items (2240 ms) ($F_1(1,7) = 4.29, p = .08; F_2(1, 38) = 4.97, p = .03$), which as noted in Experiment 5 might have been due to sensitivity to animacy information that differed between the verb bias conditions. The normal controls also showed a larger ambiguity effect for the equi-biased (191 ms) vs. transitive-biased verb items (-40 ms) (vs. 190 ms and 94 ms effect in the younger comprehenders, respectively). However, similar to the results of the younger comprehenders, neither the interaction between Verb Bias and Ambiguity nor the main effect of Ambiguity was significant for the normal controls ($F_s < 2$).

Patients

The accuracy and all reading time data of each patient were separately examined in a mixed 2 x 2 ANOVA with items serving as a random factor (Verb Bias was a between-item factor and Ambiguity was a within-item factor). A summary of the accuracy results is presented in Table 6.3 and a summary of the reading time results for the main verb (syntactic error cue) region is presented in Table 6.4.

Grammaticality judgment accuracy

Table 6.3. Experiment 6 – garden path grammaticality judgment: Mean accuracy (and standard errors) by condition for normal controls and patients, and *F*-statistics from item analysis for patients.

Condition	Controls	LC	DW	ML	($F_{1,38}$)	LC	DW	ML
EQ A	72% (11)	65% (10)	85% (9)	60% (11)	Verb Bias	1.55	<1	<1
EQ U	79% (8)	75% (9)	90% (9)	80% (10)	Ambiguity	<1	1.42	4.77*
TR A	72% (10)	80% (10)	80% (9)	55% (11)	B x A	<1	1.27	<1
TR U	73% (11)	85% (9)	70% (9)	70% (10)				

A: ambiguous U: unambiguous B x A: Verb Bias x Ambiguity * indicates $p < .05$
 EQ: equi-biased verb TR: transitive-biased verb

Patient LC. Overall, LC performed slightly above the mean of the normal control group (normal mean = 74% correct, $SD = 28$), averaging 76% correct on the experimental sentences in this task. Collapsing across verb bias conditions, he was slightly though not significantly less accurate on the ambiguous (73% correct) vs. unambiguous sentences (80% correct) (main effect of Ambiguity: $F < 1$). Similar to the normal controls, his ambiguity effect was also numerically larger for the equi-biased (10% effect) vs. transitive-biased verb items (5% effect) (Ambiguity x Verb Bias interaction: $F < 1$). Collapsing across ambiguity conditions, he was numerically less accurate on the equi-biased (70% correct) vs. transitive-biased verb items (83% correct) (main effect of Verb Bias: $F < 1.6$).

Patient DW. Overall, DW performed better than the normal controls and LC, averaging 81% correct on the experimental sentences in this task. Similar to LC, she did not show any significant main effects or a significant interaction effect in the ANOVA item analysis ($F_s < 1.5$). Collapsing across verb bias conditions, she was slightly more accurate on the ambiguous (83% correct) than on the unambiguous sentences (80% correct). She showed a 5% ambiguity effect for the equi-biased verb items but a reverse 10% ambiguity effect for the transitive-biased verb items. Collapsing across ambiguity conditions, similar to the normal controls and different from LC, she was less accurate on the transitive-biased (75% correct) than on the equi-biased verb items (88% correct).

Patient ML. ML also performed better in this experiment than in Experiment 5, although his overall accuracy still averaged under the normal mean at 66% correct. He was the only patient who showed a reliable main effect of Ambiguity in the ANOVA item analysis. He was significantly less accurate on the ambiguous (58% correct) vs. unambiguous sentences (75% correct) (main effect of Ambiguity: $F(1, 38) = 4.77, p = .04$). His ambiguity effect was slightly though not significantly larger for the equi-biased (20% effect) than for the transitive-biased verb items (15% effect) (Verb Bias x Ambiguity interaction: $F < 1$). Collapsing across ambiguity conditions, he was less accurate on the transitive-biased (63% correct) than on the equi-biased verb items (70% correct; main effect of Verb Bias: $F < 1$).

Comparison against the normal group. As in the previous experiments, the effects of each patient were compared against the normal means using the Crawford and Howell's (1998) modified t-test procedure. Performance was considered impaired if the

patient showed an effect that was significantly greater than the normal mean (significance level at $p < .05$, one-tailed, unless specified otherwise).

Compared to the overall ambiguity effect of 4.4% ($SD = 7.5$) on accuracy in the normal control group, LC's overall ambiguity effect of 7.5% and DW's overall reverse ambiguity effect of -2.5% were both within 1 standard deviation from the normal mean while ML's overall ambiguity effect of 17.5% was 1.74 standard deviations above and marginally significantly greater than the normal mean ($t(7) = 1.64, p = .07$). Within the equi-biased verb condition, all of the patient's ambiguity effects were within 1 standard deviation from the normal mean (7.5% effect, $SD = 14.1$) and not significantly greater than the normal mean (LC showed a 10% effect, DW 5% effect, and ML 20% effect; all $ts < 1$). Within the transitive-biased verb condition, LC's ambiguity effect of 5% was within 1 standard deviation from the normal mean (1.3% effect, $SD = 5.2$) and not significantly greater than the normal mean ($t < 1$). DW's reverse ambiguity effect of 10% for the transitive-biased verb items was 2.17 standard deviations below the normal mean and not significantly different from the normal mean ($t(7) = -2.05, p = .08$, two-tailed). In contrast to the other two patients, ML's ambiguity effect of 15% for transitive-biased verbs was 2.66 standard deviations above and significantly greater than the normal mean ($t(7) = 2.50, p = .02$). Finally, compared to the overall verb bias effect in the normal group (mean = 3.1%, $SD = 7.5$), LC's reverse -12.5% effect was 2.08 standard deviation below but not significantly smaller than the normal mean ($t(7) = -1.96, p = .09$, two-tailed), while DW's 12.5% effect and ML's 7.5% effect were 1.25 standard deviations and .58 standard deviation above the normal mean, respectively, which were also not significantly greater than the normal mean ($t(7) = 1.17, p = .14$, and $t < 1$, respectively).

Further verification from reading time at the syntactic cue – main verb region

As in previous RT analyses, both untransformed and log-transformed reading time data were analyzed for the patients (see Verhaeghen & De Meersman, 1998). In the untransformed reading time analysis, data points that were 2.5 *SD* beyond the mean of each condition in each participant were replaced with the 2.5 *SD* cutoff value (less than 4% of each patient's data were affected). All data (untrimmed) were included in the log-transformed reading time analysis. As in Experiment 5, a patient's RT performance was considered impaired if s/he showed an ambiguity effect that was significantly above or below the normal mean (significance level at $p < .05$, two-tailed, according to the Crawford & Howell's 1998 modified t-test procedure).

Table 6.4. Experiment 6 – garden path grammaticality judgment: Mean reading times (and standard errors) by condition at the main verb region for normal controls and patients, and *F*-statistics from item analysis for patients.

Cond	Controls	LC	DW	ML	F(1,38)	LC	DW	ML
EQ A	1168 (156)	1058 (92)	923 (51)	2828 (397)	Verb Bias	<1	<1	<1
EQ U	921 (114)	811 (45)	851 (39)	2356 (293)	Ambiguity	15.11**	6.74*	4.30*
TR A	1144 (143)	1098 (92)	913 (51)	3057 (397)	VB x A	<1	<1	<1
TR U	928 (121)	837 (45)	809 (39)	1973 (293)				

EQ: equi-biased verb TR: transitive-biased verb VB x A: Verb Bias x Ambiguity
A: ambiguous U: unambiguous * indicates $p < .05$ ** indicates $p < .01$

In striking contrast to Experiment 5, the ANOVA item analysis showed that both DW and ML were reliably sensitive to the syntactic error cue at the main verb in this grammar-focused task (main effect of Ambiguity: $F(1, 38) = 6.74$ and 4.30 , respectively, both $ps < .05$). LC, as in Experiment 5, also showed a reliable main effect of Ambiguity at the main verb ($F(1, 38) = 15.11$, $p < .001$). No other effects, main effect of Verb Bias or the interaction between Ambiguity and Verb Bias, were significant for any of the patients (all $Fs < 1$)

Compared to the normal group, who showed an overall ambiguity effect of 232 ms ($SD = 156$), LC's overall ambiguity effect of 254 ms and DW's overall ambiguity effect of 88 ms were within 1 standard deviation away from the normal mean and not significantly different from the normal mean ($-1 < \text{all } ts < 1$ for both untransformed and log-transformed data), whereas ML's overall ambiguity effect of 778 ms was 3.49 standard deviations above and significantly greater than the normal mean in the untransformed data ($t(7) = 3.29, p = .01$), though the effect was not significantly greater than the normal mean after log-transformation ($t < 1$). Within the more interfering transitive-biased verb condition, LC's ambiguity effect of 260 ms and DW's ambiguity effect of 103 ms were within 1 standard deviation away from the normal mean (217 ms, $SD = 147$) and not significantly different from the normal mean ($-1 < \text{all } ts < 1$), while ML's ambiguity effect of 1084 ms was 5.89 standard deviations above the normal mean and significantly greater than the normal mean in the untransformed data ($t(7) = 5.56, p < .001$; after log-transformed data: $t < 1$). Within the equi-biased verb condition, all three patient's ambiguity effects were within 1.5 standard deviations from the normal mean and not significantly different from the normal mean ($-1 < ts < 1.2$).

Discussion of Experiment 6

Consistent with prior research that shows a dissociation between comprehension and grammaticality judgment performance (e.g., Linebarger et al., 1983), the patients performed much better on this grammaticality judgment task than on the comprehension task in Experiment 5. Both the control patient LC and the LIFG patient DW showed similar effects to the normal controls in all the conditions. Only the LIFG patient ML showed an exaggerated ambiguity effect in the more demanding transitive-biased verb

condition. Recall that ML had a slightly worse interference resolution deficit than DW on the executive control tasks (Experiment 4) and on the semantic reinterpretation task (Experiment 5). Taken at face value, the results suggest that executive control has a role, albeit a much more limited role, in the reinterpretation of syntactic structures. The current results are clearly not consistent with the general syntactic deficit hypothesis, as DW showed no deficit on grammaticality judgments and ML showed a deficit that was specific to the more interfering transitive-bias verb condition.

A comparison between the results of the normal (older) controls in this experiment and those of the younger comprehenders in Experiment 3 showed a substantially better performance of the normal controls in the ambiguous conditions (and a slightly better performance in the unambiguous conditions). This raised a concern about the repeated use of the materials across Experiment 5 and this experiment. Upon starting the task in this experiment, the participants would have seen the same experimental sentences twice (once in each ambiguity condition). It is possible that the good performance of the normal controls and the improved performance of the patients were due to strategic factors. To address that concern, a fresh group of older healthy participants ($n = 6$) was recruited to perform the task. A combined summary showing the per-condition mean accuracy for the younger healthy comprehenders in Experiment 3 (the implausible results only) and the two groups of normal (older) controls and the patients in this experiment is given in Table 6.5a. The inferential statistics for the data of the new normal controls is presented in Table 6.5b.

Table 6.5a. Experiment 6 – garden path grammaticality judgment: Mean accuracy for younger comprehenders in Experiment 3 (implausible conditions only), for returning vs. fresh normal (older) controls, and for patients.

	Equi-biased verb			Transitive-biased verb		
	Am	Unam	Effect	Am	Unam	Effect
Younger comprehenders	53%	68%	15%	48%	68%	19%
Normal controls - Returning*	72%	79%	8%	72%	73%	1%
Normal controls - Fresh*	45%	66%	21%	45%	71%	26%
Control patient LC	65%	75%	10%	80%	85%	5%
Executive control patient DW	85%	90%	5%	80%	70%	-10%
Executive control patient ML	60%	80%	20%	55%	70%	15%
Am: ambiguous	Unam: unambiguous			Effect: unambiguous minus ambiguous		

Table 6.5b. Experiment 6 – garden path grammaticality judgment: Mean accuracy (and standard errors) by condition and *F*-statistics for fresh group of normal controls.

Condition	Accuracy	Variable	F _{1,5}	F _{2,138}
EQ ambiguous	45% (8%)	Verb Bias	<1	<1
EQ unambiguous	66% (9%)	Ambiguity	13.51**	37.86**
TR ambiguous	45% (9%)	Verb Bias x Ambiguity	2.14	<1
TR unambiguous	71% (7%)			
EQ: equi-biased verb TR: transitive-biased verb ** indicates $p < .01$				

As is clear from the table, the performance of the fresh group of normal (older) participants was more similar to that of the younger comprehenders than to that of their returning age group, suggesting that the concern about inflated level of performance in this experiment is legitimate. Note that even so, patient ML still performed poorly on the ambiguous conditions and showed an exaggerated ambiguity effect in the more demanding transitive-biased verb condition. It is possible that future studies will reveal a greater degree of syntactic reinterpretation deficit in LIFG patients (under condition of strong interference from irrelevant cues) than that found in this experiment (given that they were not tested first on similar materials in a comprehension test).

General Discussion of Experiment 4 – 6

This set of experiments provides further evidence on the importance of LIFG-based executive control in correcting for misinterpretations. Experiment 5 suggests a strong involvement of executive control in semantic reinterpretation. When the sentences contained strong cues that jointly pointed toward a semantic interpretation, the LIFG patients made the interpretation as normals did. However, unlike the normals and the control patient, when that interpretation was contradicted by another semantic cue, the executive control patients were unable to revise the initial semantic interpretation, even despite the presence of an unambiguous structure marker (the comma). Experiment 6 suggests a smaller role of executive control in syntactic reinterpretation. Only the patient with the most severe executive control deficit showed an impaired performance. However, given the methodological issue in that experiment, conclusions on the degree of executive control's involvement in semantic vs. syntactic reinterpretation must await further evidence.

Alternative explanations that focus on a general level of impairments among the three patients (ML having the most severe deficit, DW next most, and LC the least) cannot fully account for the whole pattern of results. Although ML showed the most severe deficits overall in task performance, he did not show such a degree of deficit across the board. For example, in the equi-biased verb condition, which was hypothesized to induce less demand on executive control, he and LC both averaged at 70% correct on grammaticality judgments. Similarly, DW performed at a comparable level to LC in the equi-biased verb condition of the garden path comprehension task (see Experiment 5). She also performed as well as or even better than LC on the garden path grammaticality

judgment task. The results, therefore, cannot be fully explained by alternative explanations that are based on a general level of impairments among the patients.

Chapter 4.

Summary of Findings and General Discussion

The goal of this thesis was to investigate the role of executive control in the reinterpretation of the meaning and structure of garden path sentences. Reinterpreting garden path sentences is difficult because these sentences contain cues occurring early in the sentence that support an incorrect interpretation, which may proactively interfere with the search for an alternative interpretation during reinterpretation. Comprehenders who are better able at resolving interference in working memory are therefore better equipped at reinterpretation. Conversely, those who are more susceptible to proactive interference are more likely to need to take more time revising and may eventually fail to revise. The executive control hypothesis was tested on healthy younger individuals who had varying ability in resolving interference in working memory and on patients who had a deficit in resolving interference in working memory.

Main findings of the thesis

- (1) Healthy younger comprehenders who took more time to resolve interference from irrelevant verbal representations in an independent verbal executive control task took significantly more time to revise the initial meaning of object/subject garden path sentences in one of the two most interfering conditions. In the condition hypothesized to induce the strongest interference from prior misleading cues, those who were more susceptible to interference from working memory were more likely to give up on reinterpretation.
- (2) Healthy younger comprehenders who made more errors under condition of interference from irrelevant verbal representations in an independent verbal executive

control task were less successful at revising the initial structure of the garden path sentences.

- (3) Two LIFG patients with impaired executive control were unable to revise the initial semantic interpretation even despite presence of an unambiguous structural marker (the comma) supporting the correct interpretation. One patient showed normal performance in the less interfering condition while the other patient showed poor performance across the two conditions with varying degree of interference. Their deficit was in contrast to that of a control patient who benefited from the structural cue in both the interfering conditions.
- (4) Between the two LIFG patients with impaired executive control, only the patient with the more severe deficit failed at syntactic revision in the more interfering condition. The other patient and the control patient did not show any deficit when compared to the normal older controls on this task.

Taken together, the results suggest that recovery from garden path misinterpretations is not automatic. We indeed vary in our ability to recover from misinterpretations, such that those with better executive control are better at recovering from misinterpretations while those with poorer executive control are worse at it. Furthermore, the executive control mechanisms important for recovering from misinterpretations are localizable to at least one specific area in the brain – the LIFG area. Finally, and importantly, these findings are true for both semantic *and* syntactic recovery from misinterpretations.

General discussion

Semantic reinterpretation of garden path sentences appears to depend upon several major factors: (a) availability of semantic error cues, (b) amount of interference

from previous material, and (c) ability to resolve interference in working memory. When a semantic error cue is available, most healthy younger and older comprehenders were able to abandon the initial interpretation in both the less and more interfering conditions (as shown by their high accuracy levels, averaging almost 90%, across the two verb bias conditions). Without the patient's results, the lack of a significant correlation between the verbal executive control and the semantic recovery measure in the healthy younger comprehenders for the conditions in which a semantic error cue was available (Experiment 1, implausible conditions) would suggest that semantic reinterpretation proceeds automatically once successfully triggered. However, the results from the patients, who were tested on the implausible sentences only, show that even in those conditions semantic reinterpretation is not automatic. In presence of a semantic error cue, the LIFG patient DW showed a clear pattern of semantic reinterpretation performance that was related to the amount of interference induced from previous segments of the material. She performed as well as the control patient LC when the ambiguous verb induced a minimal amount of interference (equi-biased verbs). Her performance dramatically dropped when the ambiguous verb was more interfering (transitive-biased verbs). In that condition, she stuck with the initial interpretation and endorsed it in about two-thirds of her responses (collapsed across ambiguity conditions). The other LIFG patient ML showed a different pattern of response. Unable to resolve the conflict between the current evidence (e.g., *the child was missing the uncle at home*) and the initial interpretation (e.g., *the uncle visited the child*), he endorsed one of the two pieces of information one half of the time. Even when the semantic error cue was available and the verb was less interfering, overriding the initial interpretation was still difficult for him.

The patient's results may appear to suggest that syntactic reinterpretation of the object/subject garden path sentences depends upon intact phonological short-term memory capacity. A major difference between ML – who showed a syntactic reinterpretation deficit in Experiment 6 and DW – who did not show such a deficit, despite the fact that both had an interference resolution deficit, is that ML had a much more restricted phonological short-term memory capacity than DW (phonological short-term memory composite score = -1.74 for ML vs. .28 for DW). Intact phonological short-term memory has previously been suggested to be important for reaccessing the word form of the ambiguous verb in garden path sentences (e.g., the subordinate verb), which then allows for the correct alternative structure associated with the verb to gain entry into working memory (e.g., Friedmann & Gvion, 2007). According to this account, ML's severe phonological short-term memory deficit caused him to be unable to reaccess the word form of the ambiguous verb, eventually failing to revise the initial structure of the garden path sentence.

The phonologically based explanation above fails to account for several aspects of the results. First, ML had shown *exaggerated phonological* interference effects in working memory (Experiment 4), suggesting that he might have had phonological representations from the previous segment of the material, including that of the subordinate verb, available at reinterpretation. Second, ML had a more severe phonological short-term memory deficit than LC (LC's phonological short-term memory composite score = .11 vs. -1.74 in ML), yet, he performed as well as LC in the other ambiguous condition (ambiguous, equi-biased verb condition: 60% correct vs. LC = 65% correct). Third, both ML and DW had a semantic reinterpretation deficit despite their

differences in phonological short-term memory capacity. One may argue that lexical reaccess is critical (or more critical) for syntactic reinterpretation than for semantic reinterpretation, but such explanations need further assumptions, hence not parsimonious either way.

DW's dramatically improved performance from the garden path comprehension to the garden path grammaticality judgment task needs to be accounted for, as does ML's. Perhaps the task context had an asymmetrically large effect on the patient's vs. normal older comprehender's performance. Given their restricted semantic short-term memory capacity, the patients might have capitalized on the instructions and strictly focused on grammatical aspects of the sentence. By constraining the search space to syntax-related representations, activation from other information, especially from sources of semantic information, will decrease and not interfere as much. Because of the reduced amount of interference in working memory, the patients may as a result perform better than they normally do.

Limitations of the study

The current study used a modified nonverbal version of the Stroop task to tap nonverbal executive control. However, the version as used in the study was relatively easy compared to the verbal version (as shown by perfect accuracy and shorter mean response latency compared to the verbal version). As a result, the relationship between nonverbal executive control and garden path reinterpretation could not be assessed on the accuracy measure. That does not allow for strong conclusions to be drawn regarding the role of domain-specific vs. domain-general executive control in garden path

reinterpretation. A different task should be used in future studies to tap nonverbal executive control (e.g., the antisaccade task, Miyake et al., 2000).

Use of the comma cue to mark intransitivity in the control condition of the garden path processing tasks has obvious advantages as well as certain disadvantages. Other possible control versions for the object/subject garden path sentences include reversing the order of the clauses (e.g., *The woman attended the party by herself while the man coached*) and inserting an additional noun phrase to serve as the object of the subordinate verb (e.g., *While the man coached the student the woman attended the party by herself*). However, the comma cue is preferred because it allows investigators to present virtually the same garden path sentence (plus the comma) as control. But as the results have shown, the cue did not cleanly eliminate consideration of the object interpretation in the control condition. This might have led to some loss of power in detecting the experimental effects. Other garden path structures with a clearly unambiguous structure, such as the object/sentential complement structure (e.g., *Mary confirmed (that) the trip would be delayed indefinitely*), could be used to circumvent the problem.

Third, and most seriously, the within-subject design used across the garden path experiments with the patients invites alternative explanations of the results that cannot be fully addressed without further experimentation. New patients should be tested to address the issues.

Future directions

In addition to semantic reinterpretation, the current work provides encouraging evidence that executive control is involved in syntactic reinterpretation. As grammatical

knowledge is widely assumed to be inaccessible to consciousness, one may ask what types of information are consulted and how the different sources of information are weighted while the comprehender is searching for an alternative structural representation of the material. More pressingly, it is unclear how the central executive “knows” what is relevant to attend and what it is not relevant to ignore during the reinterpretation process. Presumably the comprehender’s metalinguistic knowledge will have some role (constructive or not) in the process. Other factors such as the task context and the informativeness of the error cues (e.g., Fodor & Inoue, 2000) may all play a role in guiding the process.

References

- Adams, B. C., Clifton Jr, C., & Mitchell, D. C. (1998). Lexical guidance in sentence processing? *Psychonomic Bulletin & Review*, 5(2), 265-270.
- Baddeley, A.D. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A. (1998). The central executive: A concept and some misconceptions. *Journal of the International Neuropsychological Society*, 4(05), 523-526.
- Baddeley, A. D., & Hitch, G. J. (1994). Developments in the Concept of Working Memory. *Neuropsychology*, 8, 485-485.
- Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple component model. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 28-61). New York: Cambridge University Press.
- Barde, L. H. F., Schwartz, M. F., & Thompson-Schill, S. (2006). The role of left inferior frontal gyrus (LIFG) in semantic short-term memory: A comparison of two case studies. *Brain and Language*, 99(1-2), 71-72.
- Berndt, R. S., & Caramazza, A. (1980). A redefinition of the syndrome of Broca's aphasia: Implications for a neuropsychological model of language. *Applied Psycholinguistics*, 1(3), 225-278.
- Bever, T. G. (1970). The cognitive basis for linguistic structures. *Cognition and the development of language*, 279, 353.
- Boland, J. E., & Cutler, A. (1996). Interaction with autonomy: Multiple output models and the inadequacy of the Great Divide. *Cognition*, 58(3), 309-320.
- Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, 42(4), 368-407.

- Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist (Neuropsychology Development and Cognition Section D)*, 12(4), 482-486.
- Dick, F., Bates, E., Wulfeck, B., Utman, J. A., Dronkers, N., & Gernsbacher, M. A. (2001). Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review*, 108(4), 759-788.
- Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, 92(1-2), 145-177.
- Dunn, L. M., & Dunn, L. M. (1981). *Peabody picture vocabulary test-revised*. Circle Pines, MN: American Guidance Service.
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47(2), 164-203.
- Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25(3), 348-368.
- Ferreira, F., & Henderson, J. M. (1990). Use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(4), 555-568.
- Ferreira, F., & Henderson, J. M. (1991). How is verb information used during syntactic parsing? In G. B. Simpson (Ed.), *Understanding Word and Sentence* (pp. 305-330). Amsterdam: North-Holland.
- Fodor, J. D., & Inoue, A. (2000). Syntactic features in reanalysis: Positive and negative

- symptoms. *Journal of Psycholinguistic Research*, 29, 25–36.
- Frazier, L., & Clifton, C. (1998). Sentence reanalysis and visibility. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 143–176). Dordrecht: Kluwer Academic Publishers.
- Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6(4), 291-325.
- Freedman, M. L., & Martin, R. C. (2001). Dissociable components of short-term memory and their relation to long-term learning. *Cognitive Neuropsychology*, 18(3), 193-226.
- Friedmann, N., and Gvion, A. (2007). As Far as Individuals with Conduction Aphasia Understood these Sentences were Ungrammatical: The Garden Path in Conduction Aphasia. *Aphasiology*, 21, 570-586.
- Gahl, S., Jurafsky, D., & Roland, D. (2004). Verb subcategorization frequencies: American English corpus data, methodological studies, and cross-corpus comparisons. *Behavior Research Methods, Instruments, & Computers*, 36(3), 432-443.
- Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, 37(1), 58-93.
- Gorrell, P. (1995). *Syntax and Parsing*. Cambridge University Press.
- Grodzinsky, Y. (1990). *Theoretical perspectives on language deficits*. The MIT Press.
- Grodzinsky, Y. (1995). A restrictive theory of agrammatic comprehension. *Brain and Language*, 50(1), 27-51.

- Grodzinsky, Y., & Friederici, A. D. (2006). Neuroimaging of syntax and syntactic processing. *Current Opinion in Neurobiology*, 16(2), 240-246.
- Hamilton, A. C., & Martin, R. C. (2005). Dissociations among tasks involving inhibition: A single-case study. *Cognitive, Affective, & Behavioral Neuroscience*, 5(1), 1-13.
- Hamilton, A. C., & Martin, R. C. (2007). Proactive interference in a semantic short-term memory deficit: Role of semantic and phonological relatedness. *Cortex*, 43(1), 112-123.
- Hasher, L., Lustig, C., & Zacks, R. (2007). Inhibitory mechanisms and the control of attention. *Variation in working memory*, 227-249.
- Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. *The psychology of learning and motivation: Advances in research and theory*, 22, 193-225.
- January, D., Trueswell, J. C., & Thompson-Schill, S. L. (2009). Co-localization of Stroop and Syntactic Ambiguity Resolution in Broca's Area: Implications for the Neural Basis of Sentence Processing. *Journal of Cognitive Neuroscience*, 21(12), 2434-2444.
- Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience*, 139(1), 181-193.
- Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proceedings of the National Academy of Sciences*, 95(14), 8410.
- Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, 52(2),

205-225.

Lewis, R. L. (1993). *An architecturally-based theory of human sentence comprehension*.

Unpublished doctoral dissertation, Carnegie Mellon University.

Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition*, 13(3), 361-392.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676-703.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological bulletin*, 109(2), 163-203.

Martin, R. C. (2006). The neuropsychology of sentence processing: Where do we stand? *Cognitive Neuropsychology*, 23(1), 74-95.

Martin, R. C., & He, T. (2004). Semantic short-term memory and its role in sentence processing: A replication. *Brain and Language*, 89(1), 76-82.

Martin, R. C., & Lesch, M. F. (1996). Associations and dissociations between language impairment and list recall: Implications for models of STM. *Models of short-term memory*, 149-178.

Martin, R. C., Lesch, M. F., & Bartha, M. C. (1999). Independence of input and output phonology in word processing and short-term memory. *Journal of Memory and Language*, 41(1), 3-29.

Mason, R. A., Just, M. A., Keller, T. A., & Carpenter, P. A. (2003). Ambiguity in the brain: What brain imaging reveals about the processing of syntactically ambiguous sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1319-1338.

- Mendelsohn, A. (2002). Individual differences in ambiguity resolution: Working memory and inhibition. *Unpublished doctoral dissertation, Northeastern University.*
- Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., et al. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cognitive Brain Research*, *12*(3), 467-473.
- Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. *Attention and performance XII*, 601–618.
- Mitchell, D. C., Cuetos, F., Corley, M. M. B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research*, *24*(6), 469-488.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100.
- Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(3), 263-281.
- Novick, J.M., Kan, I.P., Trueswell, J.C., & Thompson-Schill, S.L. (2009). A case for conflict across multiple domains: Memory and language impairments follow damage to ventrolateral prefrontal cortex. *Cognitive Neuropsychology*, *26*(6), 527-567.
- Patson, N. D., Darowski, E. S., Moon, N., & Ferreira, F. (2009). Lingering

- misinterpretations in garden-path sentences: Evidence from a paraphrasing task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(1), 280-285.
- Pickering, M. J., Traxler, M. J., & Crocker, M. W. (2000). Ambiguity resolution in sentence processing: Evidence against frequency-based accounts. *Journal of Memory and Language*, 43(3), 447-475.
- Pritchett, B. (1988). Garden Path Phenomena and the Grammatical Basis of Language Processing, *Language*, 64, 539--576.
- Pritchett, B. L. (1992). *Grammatical competence and parsing performance*. Chicago, IL: University of Chicago Press.
- Roach, A., Schwartz, M. F., Martin, N., Grewal, R. S., & Brecher, A. (1996). The Philadelphia naming test: Scoring and rationale. *Clinical Aphasiology*, 24, 121-134.
- Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, 29, 86-102.
- Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology-Human Perception and Performance*, 22(2), 461-479.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643-662.
- Sturt, P. (1995). Incorporating "unconscious reanalysis" into an incremental, monotonic

- parser. In *Proceedings of the seventh conference on European chapter of the Association for Computational Linguistics* (pp. 291-296). Morgan Kaufmann Publishers Inc. San Francisco, CA, USA.
- Sturt, P. (2007). Semantic re-interpretation and garden path recovery. *Cognition*, 105(2), 477-488.
- Sturt, P., Pickering, M. J., & Crocker, M. W. (1999). Structural change and reanalysis difficulty in language comprehension. *Journal of Memory and Language*, 40(1), 136-150.
- Thompson-Schill, S. L., Jonides, J., Marshuetz, C., Smith, E. E., D'Esposito, M., Kan, I. P., et al. (2002). Effects of frontal lobe damage on interference effects in working memory. *Cogn Affect Behav Neurosci*, 2(2), 109-120.
- Traxler, M. J. (2005). Plausibility and verb subcategorization in temporarily ambiguous sentences: Evidence from self-paced reading. *Journal of psycholinguistic research*, 34(1), 1-30.
- Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language*, 35(4), 566-585.
- Trueswell, J. C., & Kim, A. E. (1998). How to prune a garden path by nipping it in the bud: Fast priming of verb argument structure. *Journal of Memory and Language*, 39(1), 102-123.
- Trueswell, J. C., & Tanenhaus, M. K. (1994). Toward a lexical framework of constraint-based syntactic ambiguity resolution. *Perspectives on sentence processing*, 155-179.
- Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in

sentence processing: Separating effects of lexical preference from garden-paths.

Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 528-528.

Van Gompel, R.P.G., & Pickering, M.J. (2001). Lexical guidance in sentence processing:

A note on Adams, Clifton, and Mitchell (1998). *Psychonomic Bulletin and Review*, 8, 851-857.

Verhaeghen, P., & Cerella, J. (2002). Aging, executive control, and attention: A review

of meta-analyses. *Neuroscience and Biobehavioral Reviews*, 26(7), 849-857.

Verhaeghen, P., & De Meersman, L. (1998). Aging and the Stroop effect: A meta-

analysis. *Psychology and Aging*, 13, 120-126.

Wilson, M. P., & Garnsey, S. M. (2009). Making simple sentences hard: Verb bias effects

in simple direct object sentences. *Journal of Memory and Language*, 60(3), 368-392.

Appendix A. Garden path sentences used in Experiment 1-3 and 4-5

Experiment 1-3 included both the plausible and implausible version (the implausible version is given in parenthesis). The slashes indicate regions of presentation. Experiment 5-6 included only the implausible version. Unambiguous control sentences were created by adding a comma after each subordinate verb.

Equi-biased verbs

While the general / advanced / the troops / moved / quickly in good order (back home early instead).

When the storm / broke / the windows / were / damaged beyond any repair (safe from any damage).

While the man / coached / the woman / attended / to the helpful advice (the party by herself).

While the thug / confessed / the crime / was / sounding like self defense (occurring outside the church).

When the robots / cracked / the problem / was / quite easy to solve (more difficult to solve).

While the jury / decided / the case / was / sealed from the public (never considered by them).

While the substance / dissolved / the stain / became / smaller and less visible (bigger and more visible).

While the husband / drove / the minivan / was / kept under good control (towed to the mechanic).

While the contestants / guessed / the words / were / difficult but still retrievable (always available as cues).

While the group / hunted / the deer / ran / quickly into the forest (joyfully around the cage).

When the boy / passed / the ball / stayed / low inside the court (securely in his hands).

While the workers / protested / the decision / was / reconsidered by the board (still not reached yet).

When the woman / refused / the man / became / pessimistic about his chances (the next choice naturally).

While the butter / softened / the dough / became / more moist than before (much dryer and harder).

When the couple / split / the bills / were / miscalculated again and again (paid completely by him).

While the racer / started / the Ferrari / went / noisy for some time (past him and others).

While the gangster / swore / the oath / was / made briefly and solemnly (created by fellow gangsters).

While the wife / swept / the porch / was / getting cleaner and cleaner (being built right outside).

While the boy / swung / the bat / stayed / in the zone awhile (with the friend instead).

While the novelist / wrote / the note / was / elaborated in a precise manner (delivered to him anonymously).

Transitive-biased verbs

While the lion / attacked / the zebra / was / fearful and yet unyielding (distant from the scene).

While the actress / called / the manager / tried / to respond with patience (to eavesdrop on her).

While the uncle / carved / the turkey / was / sliced with great precision (roasted in the oven).

When the gambler / cheated / the friend / gave / him a good beating (him lots of help).

While the aunt / cleaned / the plates / were / all over the sink (all in the cabinet).

While the mother / cooked / the chicken / sent / out a tasty aroma (its mate warning signals).

While the father / drew / the daughter / stayed / still on the grass (out with her friend).

While the housemaid / dusted / the table / collapsed / loudly onto her feet (loudly in another room).

While the fisherman / ate / the catfish / got / stuck in his throat (out of the tank).

While the host / entertained / the relative / was / fond of the jokes (alone with the dog).

When the grandmother / forgot / the grandson / was / a stranger to her (still remembered by her).

While the king / governed / the country / was / keen on welcoming foreigners (keen on invading them).

While the horses / kicked / the farmer / stayed / calm and in control (relaxed in the sauna).

While the assassin / killed / the spy / was / resisting until the end (outside the door watching).

When the girl / left / the boy / was / sad for many months (also away with her).

While the lady / performed / the ceremony / went / along without a problem (on without her presence).

When the judge / read / the verdict / was / carefully analyzed by him (reached in another room).

While the student / sketched / the girl / was / focused on her poses (at the interview alone).

While the assistant / typed / the form / was / made clear and exact (sent to her mistakenly).

While the uncle / visited / the child / was / acting nice and quiet (missing him at home).

Appendix B
Examples of filler sentences used in Experiment 1-2 and 5 (comprehension task)

While the vendor / advertised / the product / looked / more appealing than normal.
 Did the product look very appealing normally?

While the rabbit / amused / the father / the hamster / worried / him quite a bit.
 Did the hamster amuse the father?

When the twister / approached, / the farmers / were / working in the field.
 Were the farmers in danger?

While the candidates / debated / the plan / the moderator / tried / to stop them in vain.
 Did the candidates agree on the plan?

While the athlete / denied / the charges / the committee / questioned / his honesty.
 Did the athlete deny the charges?

While the uncle / fastened / the seat belt / the plane / landed / onto the exotic island.
 Did the uncle go to an island?

While the kite / flew, / the girl / imagined / herself riding away on it.
 Did the girl want to ride the kite?

While the carpenter / polished / the wood / got / thinner but not smoother.
 Did the wood get smoother?

While the student / returned / the books / the library / had / an emergency evacuation drill.
 Did the library have an evacuation drill?

While the girlfriend / rolled / the pastries / the boy / grilled / the vegetables.
 Did the boy roll the pastries?

While the sculptor / rotated / the vase / he / experienced / a sharp pain in the shoulder.
 Did the sculptor rotate the vase?

While the father / urged / the daughter / the son / prepared / breakfast in the kitchen.
 Did the daughter prepare breakfast?

Appendix C
Examples of filler sentences used in Experiment 3 and 6 (grammaticality judgment)

While the detective / reviewed / the evidence / the friend / brought / to him too late.

While the foundation / established / the celebrities / worked / to promote its cause.

While the scene / amused / the architect, / the comedian / hated / a great deal.

While the man / took, / his condition / improved / much beyond expectation.

While the witness / recounted / the story / mentioned / some inaccurate details.

While / supported / the plan / the vice-president / was / strongly against it.

While the husband / selected / the wine, / got / from the seafood section.

While / entered / the house, / the smell / alerted / him to the fire.

While the man / carried / the hoses / the neighbor / offered / to help him.

While the lawyer / framed / the question / the witness / got / confused by the wordings.

When the shepherd / built / the barn / the farmer / brought / him lots of woods.

While the man / signed / the petition / his partner / told / him it was invalid.